INVESTIGATIONS OF THE EFFECTS OF READING SKILL AND EPISTEMOLOGICAL BELIEFS ON LEARNING IN GENERAL CHEMISTRY, AND HOW STUDENTS USE TEXTBOOKS TO STUDY

Rene Buell
University of New Hampshire, Durham

Follow this and additional works at: https://scholars.unh.edu/dissertation

Recommended Citation
https://scholars.unh.edu/dissertation/2425

This Dissertation is brought to you for free and open access by the Student Scholarship at University of New Hampshire Scholars' Repository. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of University of New Hampshire Scholars' Repository. For more information, please contact nicole.hentz@unh.edu.
INVESTIGATIONS OF THE EFFECTS OF READING SKILL AND EPISTEMOLOGICAL BELIEFS ON LEARNING IN GENERAL CHEMISTRY, AND HOW STUDENTS USE TEXTBOOKS TO STUDY

BY

RENÉ W. BUCELL

B.S., Truman State University, 2009
M.S., Indiana University, 2013

DISSERTATION

Submitted to the University of New Hampshire in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Chemistry

September, 2018
This thesis/dissertation was examined and approved in partial fulfillment of the requirements for the degree of Ph.D. in Chemistry by:

Dissertation Director, Samuel Pazicni, Associate Professor of Chemistry

Christopher F. Bauer, Professor of Chemistry
Margaret E. Greenslade, Associate Professor of Chemistry
Erik B. Berda, Associate Professor of Chemistry
Jade Caines Lee, Assistant Professor of Education

On September 2018

Approval signatures are on file with the University of New Hampshire Graduate School.
ACKNOWLEDGMENTS

There are a number of people I would like to thank for their support and guidance. First, I would like to thank my research director and mentor, Dr. Samuel Pazicni. The training he provided will be with me the rest of my career, as well as guidance for how to think and work like a researcher. He was supportive in all my endeavors, and always had an open ear for ideas or general thoughts. I would also like to thank Dr. Christopher Bauer for the many conversations and pieces of advice he gave to me while I was at UNH. I would like to thank the Graduate School and Chemistry Department at the University of New Hampshire for funding to conduct my research and to travel to conferences to present my results.

My family was a great system of support during my time at UNH. Although they were far away, my parents, Rob and Gala Buell, and my brother, Grant Buell, were always available to hear about my work and stresses and successes. Throughout my life I have looked up to each of them and admired the qualities in them that helped me to grow as a scientist and person.

There have been a number of educators who made a lasting impact on me and my view of chemistry, including Mr. Evan Manuel, Dr. James McCormick, and Dr. Barbara Kramer. I would also like to acknowledge my master’s degree advisor, Dr. Kenneth Caulton at Indiana University, for his support and wisdom.

I would like to thank my friends and colleagues at UNH, including Stephanie, Julia, Dan, Gordy, Charlie, and Christian. The times I have had with them made my time at UNH a very positive experience that I will look back on fondly. I would also like to thank my friends Jarryd, Kate, Katie, Lindsay, Wyatt, Sarah, and Jenna for their support.
# TABLE OF CONTENTS

ACKNOWLEDGMENTS .................................................................................................................. iii

LIST OF TABLES ............................................................................................................................. vii

LIST OF FIGURES .......................................................................................................................... x

ABSTRACT ....................................................................................................................................... xi

## CHAPTER 

<table>
<thead>
<tr>
<th>INTRODUCTION</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background</td>
<td>2</td>
</tr>
<tr>
<td>Conceptual change model for science concepts</td>
<td>3</td>
</tr>
<tr>
<td>Epistemological beliefs</td>
<td>3</td>
</tr>
<tr>
<td>Text characteristics of general chemistry textbooks</td>
<td>5</td>
</tr>
<tr>
<td>Student use of textbooks</td>
<td>8</td>
</tr>
<tr>
<td>Objectives</td>
<td>10</td>
</tr>
</tbody>
</table>

1. THE EFFECT OF READING COMPREHENSION ON LEARNING GAINS FROM READING TEXT .......................................................... 11

| Methodology | 12 |
| Measures of prior knowledge of redox and bonding representation concepts | 12 |
| Measure of reading comprehension ability | 12 |
| General experiment design | 13 |
| Participants | 14 |
| Materials | 17 |
| Text passages about chemistry concepts | 15 |
| Data Analysis and Results | 19 |
| Dependent and independent variables | 20 |
| Large scale studies | 20 |
| Controlled study | 23 |
| Exploratory analyses | 27 |
| Large scale studies | 28 |
| Study A | 29 |
| Study B | 30 |
| Study C | 31 |
| Study D | 32 |
| Effect of reading comprehension when text was read | 35 |
| Meta-analysis | 35 |
| Description | 35 |
2. INVESTIGATION OF EPISTEMOLOGICAL BELIEFS OF GENERAL CHEMISTRY STUDENTS ................................................................. 53

Objectives .................................................................................. 53
Methodology ............................................................................... 55
Setting ....................................................................................... 55
Study design ............................................................................... 55
Outcome measures ...................................................................... 56
  Epistemological beliefs .............................................................. 56
  Course performance .................................................................. 57
Participants .................................................................................. 57
Data analysis and results ............................................................ 58
Dependent and independent variables........................................ 58
  Descriptive statistics for EBAPS scores ........................................ 59
  Descriptive statistics for course performance ............................... 63
Exploratory analysis .................................................................. 64
  Change in EBAPS scores ............................................................. 64
  Difference in EBAPS scores between sexes ................................. 65
  Differences between students of Lecturer A and Lecturer B ............ 69
  Correlation between EBAPS scores and course performance .......... 74
Analysis using linear regression .................................................. 75
  Predictors of EBAPS scores at the beginning of the semester ......... 75
  Predictors of EBAPS scores at the beginning of the semester ......... 78
  Effect of lecturer on Subscale 5 scores ......................................... 79
  Effect of course performance on subscale scores at the beginning of the semester ................................................................. 80
# LIST OF TABLES

Table 1.1 Populations involved in this study and the corresponding concept inventories for the study .................................................................15
Table 1.2 Demographic statistics for participants .................................................................16
Table 1.3 Coh-Metrix measures of readability dimensions for the text passages written for bonding and redox concepts .................................................................19
Table 1.4 Descriptive statistics of Gates-MacGinitie Reading Test (Comprehension) scores following removal of univariate outliers .................................................................25
Table 1.5 Descriptive statistics of Bonding Representations Inventory pre-test scores following removal of univariate outliers .................................................................25
Table 1.6 Descriptive statistics of Redox Concept Inventory pre-test scores following removal of univariate outliers .................................................................26
Table 1.7 Descriptive statistics of Bonding Representations Inventory post-test scores following removal of univariate outliers .................................................................26
Table 1.8 Descriptive statistics of Redox Concept Inventory post-test scores following removal of univariate outliers .................................................................27
Table 1.9 Statistical linear regression results for studies A, B, C, and D, with only statistically significant predictors included .................................................................34
Table 1.10 Data included in the meta-analysis and the Fixed Effects Model output results for the effects of whether text was read, reading skill level, prior knowledge, and the interaction between prior knowledge and reading text .................................................................38
Table 1.11 Independent t-test results (α = 0.05) comparing BRI post-test raw scores between those who read the bonding text and those who did not read the bonding text .................................................................40
Table 1.12 Statistical stepwise regression analysis results for predictors of BRI post-test Z-scores for controlled group .................................................................43
Table 2.1 Demographic statistics for participants .................................................................58
Table 2.2 Descriptive statistics of Epistemological Beliefs Assessment for Physical Science (EBAPS) scores at the beginning of the semester following removal of univariate outliers ......61
Table 2.3 Descriptive statistics of Epistemological Beliefs Assessment for Physical Science (EBAPS) scores at the end of the semester following removal of univariate outliers .................. 62

Table 2.4 Descriptive statistics of midterm and final grades (normalized as Z-scores) .............. 63

Table 2.5 Independent t-test results ($\alpha = 0.05$) comparing Epistemological Beliefs Assessment for Physical Science (EBAPS) scores at the beginning of the semester with scores at the end of the semester ....................................................... 66

Table 2.6 Paired samples t-test results ($\alpha = 0.05$) comparing Epistemological Beliefs Assessment for Physical Science (EBAPS) scores at the beginning of the semester with scores at the end of the semester for students who completed both assessments ........................................ 66

Table 2.7 Independent samples t-test results ($\alpha = 0.05$) comparing Epistemological Beliefs Assessment for Physical Science (EBAPS) scores as percentages of females (dummy variable = 1) and males (dummy variable = 0) at the beginning of the semester. Statistically significantly different scores ($\alpha=0.05$) are bolded ................................................................. 67

Table 2.8 Independent samples t-test results ($\alpha = 0.05$) comparing Epistemological Beliefs Assessment for Physical Science (EBAPS) scores as percentages of females (dummy variable = 1) and males (dummy variable = 0) at the end of the semester ....................................................... 68

Table 2.9 Independent samples t-test results ($\alpha = 0.05$) comparing Epistemological Beliefs Assessment for Physical Science (EBAPS) scores as percentages of students of Lecturer A (dummy variable = 1) and Lecturer B (dummy variable = 0) at the beginning of the semester. Results for students of Lecturer A are reported first, then results for students of Lecturer B are in parentheses. Statistically significantly different scores ($\alpha=0.05$) are bolded ................................................................. 71

Table 2.10 Independent samples t-test results ($\alpha = 0.05$) comparing Epistemological Beliefs Assessment for Physical Science (EBAPS) scores as percentages of students of Lecturer A (dummy variable = 1) and Lecturer B (dummy variable = 0) at the end of the semester. Results for students of Lecturer A are reported first, then results for students of Lecturer B are in parentheses ........................................................................................................ 72

Table 2.11 Paired samples t-test results ($\alpha = 0.05$) comparing Epistemological Beliefs Assessment for Physical Science (EBAPS) scores at the beginning of the semester with scores at the end of the semester for students who completed both assessments. Results for students of Lecturer A are reported first, then results for students of Lecturer B are in parentheses ................. 73

Table 2.12 Pearson correlations for midterm Z-scores and EBAPS scores at the beginning of the semester (N=245). Statistically significant correlations ($\alpha=0.05$) are bolded ............................................. 74

Table 2.13 Pearson correlations for final Z-scores and EBAPS scores at the end of the semester (N=105). Statistically significant correlations ($\alpha=0.05$) are bolded ............................................. 75
Table 2.14 Statistical regression analysis results for predictors of EBAPS total score at the beginning of the semester, variables added stepwise to the equation. (Total points for EBAPS is 120.) .................................................................77

Table 2.15 Statistical regression analysis results for predictors of EBAPS total score at the end of the semester, variables added stepwise to the equation. (Total points for EBAPS is 120.) ...........79

Table 2.16 Statistical regression analysis results for predictors of Subscale 5 at the beginning of the semester, variables added stepwise to the equation. (Total points of Subscale 5 is 20.) ........83

Table 2.17 Statistical regression analysis results for predictors of each subscale at the beginning of the semester, variables added stepwise to the equation..............................................................84

Table 3.1 Demographic statistics for participants.................................................................99
LIST OF FIGURES

Figure 0.1 Plots of language comprehension ability and the relationship with course performance for high and low prior knowledge students (Pyburn et al., 2013) .............................................................6

Figure 0.2 Average text readability measures for popular general chemistry textbooks (Pyburn & Pazicni 2014) .................................................................................................................................8

Figure 1.1 Forest plots for the four predictors included in the meta-analysis. ................................39

Figure 1.2 BRI pre-test and post-test scores for the controlled study ..................................41

Figure 3.1 Sequential mixed methods study design for the development of the textbook use survey .................................................................98

Figure 3.2 A histogram displaying frequency of use of different materials for studying ..........120
ABSTRACT

INVESTIGATIONS OF THE EFFECTS OF READING SKILL AND EPISTEMOLOGICAL BELIEFS ON LEARNING IN GENERAL CHEMISTRY, AND DEVELOPMENT OF A TEXTBOOK USE SURVEY

by

René W. Buell

University of New Hampshire, September, 2018

Understanding how student characteristics affect learning in chemistry can influence the pedagogical strategies employed by chemistry instructors. Previous studies have investigated the effects of characteristics including prior knowledge, math ability, and motivations on course performance. Student characteristics can also influence study strategies employed by students. Few studies, however, have focused on the role of language and reading comprehension skill on learning in chemistry, and fewer have investigated the levels of epistemological beliefs of general chemistry students. Three studies are presented in this dissertation. In the first study, the effects of prior knowledge and reading comprehension skill on learning from reading text about two chemistry concepts were assessed and analyzed. Linear regression analyses were utilized to establish relationships between predictors and test scores after reading to determine whether reading comprehension skill influenced learning gains after reading text. A meta-analysis of four large-scale studies showed that prior knowledge and reading comprehension correlated with post-test scores, and that an effect called Expertise Reversal may help low prior knowledge students close the post-test score gap if they read text with certain readability characteristics.
A second study examined the epistemological beliefs held by general chemistry students, and whether those beliefs grew in sophistication over one semester. It was found that, overall, students held moderately sophisticated beliefs in all five dimensions of epistemological beliefs measured. Students that performed better in the chemistry course were predicted to have slightly more sophisticated epistemological beliefs, and women were also predicted to have more sophisticated beliefs, and these beliefs did not change over the course of the semester, based on results of regression analyses. These findings show that there is room for growth of epistemological beliefs in general chemistry.

The final study presented is the development of a survey to quantitatively assess the use of, and attitudes toward, textbooks in a general chemistry course. Student responses to three iterations of the survey were used to develop the items and language used in the survey. This survey can be used by instructors and researchers to gather quantitative data about study strategies involving textbooks.
INTRODUCTION

Students enrolled in general chemistry courses at university come to the class with a wide range of experiences (Veloo et al., 2015), prior knowledge (Scofield, 1927; Ozsogomonyan & Loftus, 1979; Botch et al., 2007; Seery, 2009), motivations and attitudes (Boz et al., 2016), and learning strategies (Li et al., 2013; Uzuntiryaki-Kondakçı & Çapa-Aydin, 2013). More student characteristics are becoming a focus of research such as reading comprehension ability (Kendeou et al., 2003; O’Reilly & McNamara, 2007; Ozuru et al., 2009; Pyburn & Pazicni, 2014; Akbaşlı et al., 2016; Reed et al., 2016) and epistemological beliefs (Hammer, 1994; Neber & Schommer-Aikins, 2002; Hammer & Elby, 2003; Cano, 2005; Mazzarone & Grove, 2013; Mohamed & El-Habbal, 2013), and particularly how these characteristics affect learning in chemistry.

Factors that contribute to success in chemistry courses, or in mastery of chemical concepts, can inform pedagogical strategies in chemistry classrooms. It is important, however, that pedagogical strategies be informed not only by students with inadequacies, but by relationships between student characteristics and learning. When teaching strategies focus on just one group of students, or students with a particular cognitive deficit, performance disparities can grow between groups. Studies of gender disparities in course performance, for example, have indicated a change in a biology course structure increased course performance scores for women, but had no effect on men (Cotner & Ballen, 2017). In order to affect increased performance, attitudes, or motivations across a group of students given a particular pedagogical strategy, it is of utmost importance to understand the relationship between individual student characteristics and the outcome measure. Rather than focusing on deficits, a focus on how characteristics interact with interventions or teaching strategies can lead to optimal learning for all students.
Of particular interest in the studies presented here are student characteristics of reading comprehension skill and sophistication of epistemological beliefs. Both characteristics can influence what students choose to do while studying chemistry, or what materials they would consider studying with, or whether they believe they can learn new and difficult material at all. By understanding more about these student characteristics and their impact on learning or course performance, instructors can begin to work with these characteristics to provide a learning environment conducive to success.

Background

Conceptual change model for science concepts

Developing an understanding of concepts in general chemistry involves integrating one’s own experience and understanding of the world. Often, students come into the chemistry classroom with pre-formed conceptions about chemistry topics that may or may not align with the scientifically accepted view of that topic. Instructors have the difficult task of correcting those misconceptions. However, conceptual change is often very difficult to achieve, as students must undergo a series of discoveries in order to recognize the fault in their conceptions, according to Posner et al. (1982). First the student must experience deficiencies in his or her current conception. Students are very unlikely to change their concepts if in the past they have worked satisfactorily well. Accommodation of a new concept will only occur if the student can no longer solve problems or make sense of phenomena with their current concept. Then, a new conception must be available and intelligible to the student. The experiences or observations of the student must lend themselves to a new understanding (or new conception), and this new conception must prove to have the capacity to solve problems which arise. Finally, the new
concept should be “fruitful” in that it has the capability of leading to new insights. All these conditions taken together suggest that conceptual change is extremely difficult to achieve (Özdemir & Clark, 2007). A common method of disseminating information in general chemistry is through textbooks or reading material. If students holding misconceptions about a topic read about that topic, is it possible for them to begin to question their understanding, and then begin to construct a new conception?

Epistemological beliefs

A scientist’s beliefs about knowledge and learning shapes how he or she interprets information (Hofer & Pintrich, 1997). In science fields, experts agree that knowledge is ever-changing, with concepts constantly being updated based on the most recent findings. The epistemological beliefs (EB) held by experts must allow for expansion of knowledge and development of new understanding (Elby, 2010). Previous studies have shown that students with differing levels of sophistication of EB view the world (or classroom) differently, and will thus perform differently in the face of challenges in life or the classroom. For example, children who hold the belief that intelligence is fixed and unchanging are more likely to give up or feel helpless when presented with a difficult task (Dweck & Leggett, 1988). Epistemological beliefs may also affect text comprehension. Schommer et al. found that college students who hold the less sophisticated belief that knowledge consists of isolated facts have more difficulty comprehending mathematical texts (1992). Another study (Schommer-Aikins et al., 2005) found that middle school students who understand that learning is not quick and instinctual (students who believe that learning takes work) were more likely to believe that solving mathematical problems requires understanding and confidence. These attitudes may affect a student’s decision to turn to text or a textbook when studying.
Beliefs about knowledge and learning can be divided into subcategories. Hammer (1994) interviewed six physics students about their EB and found three dimensions emerge: beliefs about the structure of knowledge (as a collection of isolated facts or as a single coherent system), beliefs about the content of knowledge (as formulas or as underlying concepts), and beliefs about learning (by receiving information or through an active process of constructing understanding). Schommer (1993a) suggested four subcategories based on factor analysis of a questionnaire administered to 200 postsecondary students. Two of these EB dimensions were labeled consistently with Hammer’s dimensions (simple knowledge and certain knowledge), and two new dimensions were introduced as the beliefs in innate ability and quick learning. Another dimension of EB introduced by Elby and coworkers in their Epistemological Beliefs Assessment for Physical Sciences (EBAPS) was real life applicability. This dimension probes whether the student considers scientific thinking to be applicable only in restricted spheres such as the classroom or laboratory.

Baxter Magolda (2004) uses a constructivist model to characterize levels of EB (from low to high sophistication) as absolute knowing, transitional knowing, independent knowing, and contextual knowing. There is experimental support that students gradually progress from naïve epistemological beliefs to mature beliefs (Kitchener et al., 1989), but this progression has not been extensively investigated in terms of science EB, particularly quantitatively. Chapter 2 investigates the epistemological beliefs and the influence of student characteristics of first-semester general chemistry students through use of a quantitative assessment, the Epistemological Beliefs Assessment for Physical Science (Elby et al., 2006).
Text characteristics of general chemistry textbooks

Students in general chemistry classes in the university setting are diverse in terms of preparation, prior knowledge, reading skill, and epistemological beliefs. The effects that differences in preparation and prior knowledge have on course performance have been investigated. Math ability (Leopold & Edgar, 2008), SAT scores (Pickering, 1975; Spencer 2006), logical reasoning and thinking (Bunce & Hutchinson, 1993; Lewis & Lewis, 2007; Bird, 2010) have been found to have positive correlations with general chemistry course performance. These student characteristics have been investigated mainly for identifying at-risk students, and have not been used in the development of learning or study materials.

The effect of student reading skill has also been investigated in the context of overall performance in a general chemistry course (Pyburn et al., 2013). Students’ reading skill as measured by a standardized reading comprehension assessment, the Gates-MacGinitie Reading Test (GMRT), correlated with ACS exam scores and course performance, suggesting that skilled readers with low prior knowledge may be able to compensate by being able to bridge conceptual gaps using prior knowledge while comprehending new information (Figure 0.1). On the other hand, unskilled readers may not have the ability to ignore irrelevant information, and thus have a disadvantage when presented with new information with a higher cognitive load. These findings are consistent with the Structure-Building Framework as described by Gernsbacher (1991). Understanding the effect of a student’s reading ability on learning has implications in the development of study materials designed to differentially aid students.

In general chemistry classrooms, text-based resources are often used by instructors to provide students with additional information about chemistry topics, including physical and
electronic textbooks. A common textbook may be assigned to a class which contains a diverse student population, with a wide range of reading ability and prior chemistry knowledge.

Figure 0.1. Plots of language comprehension ability (SAT scores in blue, GMRT in green) and the relationship with standardized ACS exams scores (A) and course performance (B) for high (dotted line) and low (solid line) prior knowledge students (Pyburn et al., 2013).

The effectiveness of a common text resource among a diverse group of students may be limited, as student characteristics have been shown to affect comprehension of science texts (Ozuru et al., 2009). In the context of a general biology course, students with low and high prior knowledge were given reading material with high and low levels of cohesion within the text. Prior knowledge and reading ability both acted as predictors for performance on comprehension questions. Overall, prior knowledge explained a significant amount of variance, particularly on questions that required an extensive amount of integration of information. An interaction between text cohesion and reading ability was also found to be significant, though the effect size was very small ($\eta^2 = 0.03$). It was found that the interaction was significant in the performance of text-based questions on the assessment, and that the highly cohesive text material mentioned the information required for those questions more frequently than the text with low cohesion. The
results of this study suggest that the text-based material that students may use to review concepts may itself have an effect on student retention of information.

In addition, the language used in general chemistry textbooks spans a wide range of readability between different textbooks and even within the same textbook (Pyburn & Pazicni, 2014). The five most widely general chemistry text books from the year 2012 were analyzed (Brown et al., 12th ed.; Zumdahl and Zumdahl, 9th ed.; Chang, 11th ed.; Tro, 2nd ed.; and Silberberg, 6th ed.) and compared to a best selling popular novel. The five readability dimensions investigated include narrativity, syntactic simplicity, word concreteness, referential cohesion, and deep cohesion (Figure 0.2). Narrativity refers to how closely the text follows a storyline. General chemistry textbooks have low narrativity scores compared to the novel, which is unsurprising as chemistry texts are informational. Syntactic simplicity is the measure of the degree to which shorter and simpler sentence structures appear in the text. General chemistry textbooks score low on word concreteness compared to novels, as chemistry involves many abstract ideas. On average, general chemistry textbooks use longer, more complex sentences than novels, and are therefore more difficult to read.

Referential cohesion refers to how often the text uses overlapping words and ideas in order to make connections between ideas for the reader. Because students with low prior knowledge benefit from being presented with connections between concepts, a high referential cohesion score would be preferable for general chemistry texts, but the texts analyzed were found to have referential text scores ranging from approximately 53/100 to 80/100. This finding suggests that different general chemistry texts offered to students may differentially aid students depending on their level of prior knowledge. The final readability characteristic analyzed was deep cohesion, which is the extent to which the text makes causal and intentional connectives
between ideas. Again, students with a lower level of prior knowledge may need connections between ideas made explicitly for them, while students with higher prior knowledge benefit from having to make the connections themselves. Reading comprehension skill can help low prior knowledge students make those connections on their own, and text authors have a good amount of control over the cohesion in their texts. It was found that the general chemistry texts are more cohesive than the popular novel tested. The deep cohesion scores were found to not vary widely across the texts, and they scored on average 55/100, suggesting that no certain set of students is particularly aided in terms of deep cohesion.

Figure 0.2. Average text readability measures for popular general chemistry textbooks (Pyburn & Pazicni 2014).

Textbook authors do not typically actively consider the effects that readability measures, reading ability, and prior knowledge may have on student learning when composing their texts.

Student use of textbooks

Textbook materials are rich sources of information for students enrolled in undergraduate science classes. However, these materials are often expensive, so it is important to assess the
usefulness of science textbooks for students so that instructors make more informed decisions regarding textbook selection and use. Several studies have documented student self-reported study habits (Randahl, 2012; Lopez et al. 2013), and have concluded that students do report that they use the textbook as one of their main sources of information while studying mathematics or science, but no additional information is available on how the students interact with the material. A student may “use the textbook” in the form of merely reading the chapter without taking notes. Another student may choose to highlight passages in the book, and another student may only work problems at the end of the chapter. These “uses” of the book may or may not be created equal, but no study has been done to understand the details of textbook usage by student. Therefore, even though there is evidence of textbook usage, there is little insight into the meaningfulness of the textbook usage, and so faculty may have difficulty justifying the requirement that students purchase expensive and lengthy textbook to use as reference and study materials. Ronald Gillespie (1997) wrote in a commentary to the Journal of Chemical Education that, “No matter how excellent these texts appear to the instructors who choose them, they have not succeeded in interesting the vast majority of students or in providing them with an understanding of chemistry.” Understanding whether there are particular sections which do interest the students can allow for improvement in textbook material and usage.

Chemistry textbooks have been analyzed in terms of text difficulty, quality of images and diagrams, and types of practice problems (Chiappetta et al., 2013; Dávila & Talanquer, 2010; Pyburn & Pazicni, 2014; Smith & Jacobs, 2003). Students may choose to focus on particular aspects or features of the textbook when they are studying for different reasons, but thus far no work has been done to investigate the students’ reasoning for use of these features, and so chemistry instructors only suggest or assign readings based on their personal opinions or just by
guesswork. If themes emerge regarding more popular or “useful” features of the general chemistry textbook, instructors may choose to highlight those sections of the book and encourage students to read those sections, rather than giving a blanket suggestion to read the entire chapter. This may motivate students to use the textbook more since the amount of reading may be less daunting.

Objectives

1. Because student characteristics of prior knowledge and reading comprehension ability vary among college students, it is imperative to understand the effects these characteristics may have on learning gains made while students study. The goal in the first study was to examine the effects of these characteristics on learning outcomes after reading text passages. The results in Chapter 1 assess the predictive ability of prior knowledge, student comprehension ability, reading text, and any interactions between these main effects on post-test scores about chemistry concepts. A meta-analysis was also performed to determine any patterns between four independent large-scale studies.

2. Epistemological beliefs can guide how students approach studying and learning in chemistry. The results in Chapter 2 assess the sophistication of general chemistry students’ epistemological beliefs and the growth of those beliefs during a single semester.

3. Lastly, textbooks are common resource materials employed by instructors in general chemistry courses at colleges and universities, but little work has been done to see how and why students use the textbooks when studying. The results in Chapter 3 show the development of a textbook use survey that can be utilized by instructors or researchers to investigate student textbook usage and attitudes towards textbooks in general chemistry.
CHAPTER 1

THE EFFECT OF READING COMPREHENSION ON LEARNING GAINS FROM READING TEXT

The goal of understanding how student characteristics of prior knowledge and reading comprehension skill affect learning gains students can achieve by reading text-based material leads to the following research questions:

1. To what degree will reading about a general chemistry concept promote learning gains?

Because textbooks (both physical and online) are such ubiquitous reference materials in general chemistry classrooms, it is imperative to understand whether learning gains can be achieved by reading, and to what extent reading text helps a student learn general chemistry topics.

2. How does student reading comprehension influence learning gains when a text passage is used for studying a concept?

Undergraduate students have varying degrees of reading skill, and understanding how reading ability affects comprehension of text and retention of information has implications for how textbooks can be written.

3. Does an interaction between reading comprehension and prior knowledge predict learning outcomes, and to what extent?

In addition to students possessing different reading skill levels, they also have varying levels of prior knowledge. Investigating any interaction between reading ability and prior knowledge can inform textbook writers and instructors about how text can differentially aid students.
Methodology

Measures of prior knowledge of redox and bonding representation concepts

To measure the understanding of bonding representations and redox concepts, the Bonding Representations Inventory (BRI) (Luxford & Bretz, 2014) and Redox Concept Inventory (ROXCI) (Brandriet & Bretz, 2014) were completed by students at the beginning of the semester and approximately four weeks later. These concept inventories were developed based on student work and semi-structured interviews using a constructivist framework in order to probe student mastery of chemistry concepts. Such methods allow the assessments to reflect how students understand the concepts at hand, rather than how the expert developing the inventory believes students think about the concepts. Bonding concepts are often introduced during the first semester of general chemistry, and redox concepts are often introduced during the first or second semester of general chemistry.

Measure of reading comprehension ability

The standardized Gates-MacGinitie reading test (Comprehension 10/12 – Form S 4th edition) was administered to students at the beginning of the semester at the same time as the chemistry topic pre-test (MacGinitie, MacGinitie, Maria, Dreyer, & Hughes, 2000a). The test is a timed, 48-question test that measures the level of reading comprehension skill of the test-taker. It is comprised of a series of short text passages followed by four or five questions about each passage. Test results can be compared to normalized reading levels based on national (U.S.) standards. The GMRT Comprehension 10/12 Form S test was designed to be taken by high school students, and norms were published for the Fall, Winter, and Spring terms during the students’ final year. As a majority of the participants in the study were first year university students, the results of the GMRT were compared to the published Spring term norms (MacGinitie, MacGinitie, Maria, Dreyer, & Hughes, 2000b). The mean average scores of reading
comprehension ability across the five populations involved in this study were not statistically significantly different from each other, with an overall average score of 63.0%, which falls at the 49th percentile of graduating high school seniors.

General Experiment Design

A between subjects design was used in this study (Charness et al., 2012). Participants in the study completed assessments to measure their reading comprehension ability (and their level of prior knowledge for one of two general chemistry concepts at the beginning of the semester. The chemistry concepts were tested using concept inventories developed by researchers to probe the mastery of understanding of the concepts of bonding representations and of reduction-oxidation chemistry (Luxford & Bretz, 2014; Brandreit & Bretz, 2014b). Three weeks later, an intervention and post-test were administered. The intervention consisted of reading a text passage about the topic of the concept inventory the group had completed at the beginning of the semester. A short time later (between 1-3 hours), a post-test was administered, which was identical to the pre-test about the chemistry concept. Control groups of students were given unrelated texts that had similar text characteristics as the concept reading passages. Pre- and post-test scores were compared using paired samples t-tests, and linear regression was used to identify variables which were statistically significant predictors of post-test outcomes. Variables included in the analysis were pre-test scores, Gates-MacGinitie reading comprehension scores, a dummy variable representing whether the topic text was read prior to the post-test, and interactions between the main effect variables. IBM SPSS statistics software was used for the statistical analysis of results.
Participants

Participants of this study were enrolled in a traditional two-semester general chemistry course sequence and in a one-semester general chemistry course for engineers at a four-year public university with higher research activity in the northeastern United States. Five separate populations participated in the study between Fall 2014 and Fall 2015. Four experiments were conducted with over 200 students each, and the students held a range of prior knowledge of the concepts. One experiment included a tightly controlled subpopulation of low prior knowledge general chemistry students. In all experiments, students completed the GMRT and either the BRI or ROXCI at the beginning of the semester. Students in the first-semester general chemistry course completed the BRI, students in the second-semester general chemistry course completed the ROXCI, and students in the one-semester course designed for engineering majors completed either the BRI or ROXCI. The experiments for all populations were completed prior to class instruction on the topic tested. The summary of the populations in this study is shown below in Table 1.1. Demographic information of the participants was collected from school records and the results are displayed in Table 1.2. Data regarding student prior knowledge and reading comprehension ability were collected by assessments at the beginning of each semester. Post-test measurements for concept knowledge were collected approximately three weeks into the semester. Descriptive statistics for these data are discussed in the next section, and analysis of these data with respect to the research questions for this study are discussed further. Institutional Review Board (IRB) approval was sought and determined to be exempt from IRB oversight because the research design was in keeping with normal classroom practices. Copies of IRB approval letters are provided in Appendix A.
For the experiments with larger populations, the text passage was read at the beginning of their general chemistry laboratory session for the week, and the post-test was given at the end of the three-hour lab session. In the controlled experiment, the intervention and post-test were done in a conference room, with the post-test being given approximately one hour after the reading of the text passage. Control groups of students were given unrelated texts about loons or psychedelic mushrooms prior to taking the post-test.

Table 1.1. Populations involved in this study and the corresponding concept inventories for the study.

<table>
<thead>
<tr>
<th>Course</th>
<th>Semester</th>
<th>Concept inventory</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chem 403</td>
<td>Fall 2014</td>
<td>BRI</td>
<td>Controlled study</td>
</tr>
<tr>
<td>Chem 404</td>
<td>Spring 2015</td>
<td>ROXCI</td>
<td>A</td>
</tr>
<tr>
<td>Chem 405</td>
<td>Spring 2015</td>
<td>BRI</td>
<td>B</td>
</tr>
<tr>
<td>Chem 403</td>
<td>Fall 2015</td>
<td>BRI</td>
<td>C</td>
</tr>
<tr>
<td>Chem 405</td>
<td>Fall 2015</td>
<td>ROXCI</td>
<td>D</td>
</tr>
</tbody>
</table>
Table 1.2. Demographics statistics for participants.

<table>
<thead>
<tr>
<th></th>
<th>Controlled</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td>100%(^a)</td>
<td>69.2% Female 30.8% Male</td>
<td>65% Female 35% Male</td>
<td>24.7% Female 75.3% Male</td>
<td>72.5% Female 27.5% Male</td>
</tr>
<tr>
<td><strong>Class standing</strong></td>
<td>98%</td>
<td>70.3% First-years 29.7% Sophomores 4.7% Juniors 7.8% Seniors</td>
<td>55.1% First-years 32.7% Sophomores 8.3% Juniors 3.9% Seniors</td>
<td>37.0% First-years 43.1% Sophomores 6.8% Juniors 0.5% Seniors</td>
<td>67.1% First-years 24.4% Sophomores 6.8% Juniors 1.7% Seniors</td>
</tr>
<tr>
<td><strong>Academic major</strong></td>
<td>89%</td>
<td>23.1% Biology 21.5% Biomedical Science: Medical and Veterinary Sciences 12.3% Zoology 12.3% Neuroscience 6.2% Medical Laboratory Sciences 4.6% Biochemistry, Molecular and Cellular Biology 4.6% Chemistry</td>
<td>17.6% Biomedical Science: Medical and Veterinary Sciences 16.5% Biology 12.0% Zoology 6.3% Medical Laboratory Sciences 5.1% Biochemistry, Molecular and Cellular Biology 3.7% Undeclared 2.9% Chemistry</td>
<td>27.6% Mechanical Engineering 27.6% Civil Engineering 13.3% Chemical engineering 6.6% Environmental Engineering 3.9% Undeclared 3.4% Biological Engineering</td>
<td>14.3% Biomedical Science: Medical and Veterinary Sciences 12.8% Biology 10.2% Undeclared 7.3% Zoology 5.6% Biochemistry, Molecular and Cellular Biology 4.6% Medical Laboratory Sciences 1.9% Chemistry</td>
</tr>
</tbody>
</table>

\(^a\) percentage of participants for whom data are available
Materials

Text passages about chemistry concepts

Several textbooks were investigated to find short text passages about bonding and redox concepts, but none were found to have concise and centralized text about these concepts. Textbooks often introduce topics pertaining to a concept in different chapters throughout the textbook. This may make it difficult for students to make the connections between the topics unless the student has a high level of prior knowledge. The text passages used in this study were written to simulate writing found in textbooks in terms of content and text readability measures. The text passages were designed to include topics that students often hold misconceptions about in order to test whether these misconceptions can be corrected upon reading a text passage.

The text passages were written to have similar readability features as traditional text passages from textbooks, as discussed below. The concepts included in the text passages were chosen based on published literature about common misconceptions about bonding (Peterson et al., 1986; Zoller, 1990; Taber, 1997; Harrison & Treagust, 2000; Özmen, 2004) and redox concepts (Garnett & Treagust, 1992; De Jong et al., 1995; Stains & Talanquer, 2008; Österland & Ekborg, 2009; Barke, 2012; Brandriet & Bretz, 2014a). The texts were not written to address the specific misconceptions tested in the concept inventories, but rather to address misconceptions commonly found by researchers and presented in the literature. The passages can be found in Appendix B and Appendix C.

The text about bonding representations addressed several common topics where students hold misconceptions. These topics include the idea that bonds are not necessarily purely ionic or covalent, that polar covalent bonds exist because of differences in electronegativities of atoms
bound together, the tendency for metals to become cations and nonmetals to become anions by losing or gaining electrons, and that covalent compounds tend to form discrete molecules while ionic compounds form extended 3D arrays of alternating ions. The reasoning for including these ideas in the text is that it is hypothesized that if a student holds a misconception about a topic prior to reading the text passage, the student may start to overcome the misconception if confronted with the scientifically accepted understanding of the topic while reading (Nakhleh, 1992; Kendeou et al., 2003). Some of the items in the BRI address those same common misconceptions because the BRI was developed using student responses to questions about bonding.

The text about redox concepts included discussions about the difference between oxidation numbers and charge, movement of electrons from one species to another, and the definition of reductants and oxidants. These ideas were included because students have been found to hold misconceptions about these topics. The ROXCI includes some items which refer to these concepts.

The passages were limited to 600 words and the readability of the passages were analyzed using Coh-Metrix software (Graesser et al., 2004; McNamara et al., 2005). The five readability dimensions fell within a standard deviation of the average textbook readability measures found by Pyburn & Pazicni (2014) (Table 1.3).
Table 1.3 Coh-Metrix measures of readability dimensions for the text passages written for bonding and redox concepts compared to median values found for five general chemistry textbooks by Pyburn & Pazicni (2014). Percentages represent the percentile at which the measured dimension falls.

<table>
<thead>
<tr>
<th></th>
<th>Bonding representations text</th>
<th>Redox concepts text</th>
<th>Median Value for Popular general chemistry textbooks (Standard deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of words</td>
<td>456</td>
<td>605</td>
<td>-</td>
</tr>
<tr>
<td>Narrativity</td>
<td>5%</td>
<td>10%</td>
<td>22% (14)</td>
</tr>
<tr>
<td>Syntactic Simplicity</td>
<td>75%</td>
<td>49%</td>
<td>52% (21)</td>
</tr>
<tr>
<td>Word Concreteness</td>
<td>7%</td>
<td>23%</td>
<td>32% (24)</td>
</tr>
<tr>
<td>Referential Cohesion</td>
<td>86%</td>
<td>96%</td>
<td>65% (28)</td>
</tr>
<tr>
<td>Deep Cohesion</td>
<td>33%</td>
<td>83%</td>
<td>55% (26)</td>
</tr>
<tr>
<td>Flesch Kincaid Grade Level</td>
<td>14</td>
<td>13.6</td>
<td>-</td>
</tr>
</tbody>
</table>

Data Analysis and Results

All univariate outliers were eliminated from analyses using a Mahalanobis distance criterion of \( p < 0.001 \) (Tabachnick & Fidell, 2013, p 99). Five outliers were omitted from group A, no outliers were found in group B, 25 outliers were omitted in group C, and 20 outliers were omitted in group D. Two outliers were omitted from the controlled experiment. No multivariate outliers were identified. Reliability measurements for the concept inventories (as measured by Cronbach’s \( \alpha \) ) are not reported, as reliability values for concept inventories can be misleading. Students may have fragmented knowledge about concepts, and these would lead to low Cronbach’s \( \alpha \) scores (Adams & Wieman, 2011).
Dependent and independent variables

Descriptive statistics for the large scale studies

Reading comprehension

A. A total of 290 students participated in this experiment. The mean average GMRT score for this population was 29 out of 48, with a standard deviation of 9. The skewness and kurtosis scores were also within the range of ±1, and so the inferential statistics used were robust against the modest deviation from normality. The reading comprehension level was approximately a 12th grade reading level for this population as well.

B. In this experiment, 143 students participated, and they scored a mean average of 32 (±9) on the GMRT. Compared to national norms, this population scored at the 52nd percentile with a post-high school reading comprehension level. The distribution of scores was approximately normal.

C. The third large population study was completed with 396 participants. The mean average GMRT score was 30 with a standard deviation of 7, again falling in the 49th percentile according to national norms, and a 12th grade reading level. No extreme measures of skewness or kurtosis were present in this population.

D. The final larger population in this study consisted of 145 participants, who scored a mean average of 31 (±8) out of 48 points on the GMRT. These scores are very similar to the other four populations in this study, falling in the 52nd percentile of normalized scores, and corresponding to a post-high school reading comprehension level, and with skewness and kurtosis scores within ±1, so the assumption of normality applies. All the GMRT statistics are reported in Table 1.4.
**BRI pre-test scores**

Studies B and C completed the Bonding Representations Inventory in order to measure learning gains from reading text. The scores on the pre-tests for both groups had skewness and kurtosis values falling within ±1, so all inferential statistics involving the BRI pre-test values were robust against the modest difference from normality.

**B.** The mean average for the 143 participants in this population was 10 points (± 4) out of 23 points. The students enrolled in this course scored, on average, lower than students in general chemistry courses after they have received instruction about bonding concepts, but they scored higher than high school students (Luxford & Bretz, 2014). This finding is reasonable because the students enrolled in particular course generally outperform students in the traditional two-semester general chemistry sequence.

**C.** The 396 participants in this population scored a mean average value of 9 out of 23, with a standard deviation of 3 points. These results are comparable to scores found when the test was administered to high school chemistry students, 9 ± 3 (Luxford & Bretz, 2014). The summary of BRI pre-test scores are reported in Table 1.5.

**ROXCI pre-test scores**

Populations A and D completed the Redox Concepts Inventory in order to measure learning gains from reading text. The scores on the pre-tests for all three groups had skewness and kurtosis values falling within ±1, so all inferential statistics involving the ROXCI pre-test values were robust against the modest difference from normality.

**A.** A mean average score of 4 out of 18 points was scored by this population of 285 participants, with a standard deviation of 2 points. This average is lower than the average score
reported in a previous study of 5 points (± 3), but those students had been introduced to redox concepts in lecture prior to completing the inventory. The students in study A had no introduction to redox concepts in this course before the inventory was completed.

**D.** This population of 145 participants scored, on average, 4 out of 18 points (± 2) on the ROXCI pre-test. These scores were also below scores found when general chemistry students completed the test in a prior study, but the same conditions from Study A applied to Study D. The summary of ROXCI pre-test scores are reported in Table 1.6.

**BRI post-test scores**

For all populations that completed the BRI post-test, skewness and kurtosis values fell between ± 1, so the modest deviation from normality did not affect the inferential statistical analyses. Pre- and post-test scores were compared using paired samples t-test at 95% confidence. Effects of student characteristics which may lead to the change of BRI scores were investigated and are discussed in a later section. Results are presented in Table 1.7.

**B.** This population of 146 participants scored a mean average of 10 points (± 4) out of 23 on the BRI post-test, which is not statistically significantly different at 95% confidence from the pre-test of 10 ± 4 based on the paired-samples T-test (p = 0.079), with a small effect size (Cohen’s $d = 0.15$).

**C.** This BRI population of 396 participants scored a mean average of 10 points (± 4) on the BRI post-test. The difference between pre- and post-test scores was significant at 95% confidence (p < 0.001), with a small/medium effect size (Cohen’s $d = 0.38$).
ROXCI post-test scores

For both populations that completed the ROXCI post-test, skewness and kurtosis values fell between ± 1, so the modest deviation from normality did not affect the inferential statistical analyses. Pre- and post-test scores were compared using paired samples t-test at 95% confidence. Effects of student characteristics which may lead to the change of ROXCI scores were investigated and are discussed in a later section. A summary of these statistical results is provided in Table 1.8.

A. The mean average for ROXCI post-test scores was 5 (±2), which is a statistically significant increase from the pre-test score of 4 (p < 0.001), though the increase is smaller than one point out of 18. The effect size as measured by Cohen’s d is small/medium (d = 0.31).

D. The mean average the ROXCI post-test scores for this population of 145 participants was 4.5 (±1.98) out of 18 points. This was a statistically significant increase from the pre-test average score of 3.6 ±2 (p < 0.001) and a medium effect size (Cohen’s d = 0.48), though the increase of 0.9 points is small (5% of 18 possible points). Factors which may have contributed to the score increase were investigated below.

Descriptive statistics for the controlled study

Reading comprehension

Controlled Study. 459 students completed the GMRT at the beginning of the semester. The mean average score for the GMRT was 30 out of a possible 48, or 61%, which corresponds to a 12th grade reading level (46th percentile when compared to national norms from the Spring term of the final year of high school). The standard deviation was 9, or 17% of the maximum score. Based on scores from the BRI pre-test, the students who scored below average on the BRI
were invited to take the post-test, and a total of 56 students volunteered for the second phase of the study. The mean average reading comprehension score of the 56 participants was 30 out of 48 points, or 63%, which was not statistically significantly different from the scores earned by the whole population of 459 students (p = 0.415). The skewness and kurtosis values for the GMRT scores were found to be ±1, so these data are only modestly violate the assumption of normality, and so the inferential statistics employed in this study are robust (Cohen et al., 2003, p. 41).

**BRI pre-test scores**

Controlled. There were a total of 459 participants during the first phase of the experiment. These students scored an average of 9 points out of a maximum of 23 points on the BRI pre-test, with a standard deviation of 3 points. This average score corresponds with scores of high school chemistry students, which is reasonable as many students who participated in this study have come straight from high school. The students selected to participate in the study were chosen because they performed below average on the BRI pre-test. The average score of the 56 participants was 8 out of 23 points (± 2 points). These scores fall below the average scores of general chemistry and high school chemistry students (12.47 and 8.71 points out of 23, respectively) (Luxford & Bretz, 2014).

**BRI post-test scores**

Controlled. Students who completed the post-test volunteered after being invited to participate. These students scored on average a mean of 11 points (± 4) out of a possible 23 points on the post-test, an increase of 2 points, or 10.7%. This difference is statistically significant at 95% confidence (p = 0.0006), with a large effect size (Cohen’s d = 0.87). Further investigation was done to determine predictors for score change and will be discussed below.
Table 1.4 Descriptive statistics of Gates-MacGinitie Reading Test (Comprehension) scores following removal of univariate outliers.

<table>
<thead>
<tr>
<th></th>
<th>Controlled (All participants in the first phase of the study)</th>
<th>Controlled (Participants in the intervention and post-test)</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>459</td>
<td>56</td>
<td>290</td>
<td>143</td>
<td>396</td>
<td>145</td>
</tr>
<tr>
<td>Mean</td>
<td>29.5</td>
<td>30.4</td>
<td>29.4</td>
<td>31.6</td>
<td>30.1</td>
<td>30.8</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>8.5</td>
<td>8.9</td>
<td>8.5</td>
<td>8.8</td>
<td>7.5</td>
<td>8.3</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.065 (Std. Error 0.114)</td>
<td>-0.374 (Std. Error 0.319)</td>
<td>-0.058</td>
<td>-0.368</td>
<td>0.130</td>
<td>0.009</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-0.250 (Std. Error 0.227)</td>
<td>0.252 (Std. Error 0.628)</td>
<td>-0.219</td>
<td>-0.026</td>
<td>-0.691</td>
<td>-0.816</td>
</tr>
</tbody>
</table>

Table 1.5 Descriptive statistics of Bonding Representations Inventory pre-test scores following removal of univariate outliers.

<table>
<thead>
<tr>
<th></th>
<th>Controlled (All participants in the first phase of the study)</th>
<th>Controlled (Participants in the intervention and post-test)</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>459</td>
<td>56</td>
<td>143</td>
<td>396</td>
</tr>
<tr>
<td>Mean</td>
<td>8.8</td>
<td>8.2</td>
<td>9.8</td>
<td>8.8</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>3.4</td>
<td>2.5</td>
<td>3.7</td>
<td>3.0</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.706 (Std. Error 0.114)</td>
<td>-0.137 (Std. Error 0.319)</td>
<td>0.767</td>
<td>0.130</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.852 (Std. Error 0.227)</td>
<td>0.284 (Std. Error 0.628)</td>
<td>0.570</td>
<td>-0.691</td>
</tr>
</tbody>
</table>
Table 1.6 Descriptive statistics of Redox Concept Inventory pre-test scores following removal of univariate outliers.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>285</td>
<td>145</td>
</tr>
<tr>
<td>Mean</td>
<td>4.2</td>
<td>3.6</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>2.2</td>
<td>1.5</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.931 (Std. Error 0.144)</td>
<td>0.172 (Std. Error 0.201)</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>0.819 (Std. Error 0.288)</td>
<td>-0.074 (Std. Error 0.400)</td>
</tr>
</tbody>
</table>

Table 1.7 Descriptive statistics of Bonding Representations Inventory post-test scores following removal of univariate outliers.

<table>
<thead>
<tr>
<th>Controlled (Participants in the intervention and post-test)</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>56</td>
<td>143</td>
</tr>
<tr>
<td>Mean</td>
<td>11.2</td>
<td>10.4</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>4.2</td>
<td>3.8</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.047 (Std. Error 0.319)</td>
<td>0.086 (Std. Error 0.202)</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-0.822 (Std. Error 0.628)</td>
<td>-0.561 (Std. Error 0.401)</td>
</tr>
</tbody>
</table>
Table 1.8 Descriptive statistics of Redox Concept Inventory post-test scores following removal of univariate outliers.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>285</td>
<td>145</td>
</tr>
<tr>
<td>Mean</td>
<td>4.8</td>
<td>4.5</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>2.1</td>
<td>2.0</td>
</tr>
<tr>
<td>Skewness (Std. Error)</td>
<td>0.931 (0.144)</td>
<td>0.200 (0.201)</td>
</tr>
<tr>
<td>Kurtosis (Std. Error)</td>
<td>0.819 (0.279)</td>
<td>-0.165 (0.400)</td>
</tr>
</tbody>
</table>

**Exploratory analyses**

As discussed above, three research questions were explored in this study. First, the question of how reading can promote learning gains can be investigated by comparing pre- and post-test score changes of students who read a text passage about the topic tested by the concept inventories and a control group of students, who read text passages about loons or psychedelic mushrooms. Second, how reading comprehension skill level might affect learning gains. And thirdly, whether an interaction between reading comprehension and prior knowledge may moderate learning gains. First, the results of the larger scale studies (N_A = 290, N_B = 143, N_C = 396, N_D = 145) will be discussed. In all cases, GMRT scores and concept inventory scores were transformed into centralized Z-scores in order to generalize results and compare values with different scales, means, and standard deviations (Warner, 2013 p. 75). Statistical regression analysis was used to examine significant predictors for post-test scores. The variables of pre-test Z-scores, GMRT Z-scores, whether related text was read (dummy variable 1 = related text was read, 0 = text was not read), and interactions between all main effects were included in the
analyses. The variables were entered stepwise based on how much R² for the model increased by
the inclusion of the variable. Variables which increased R² the most were entered into the
regression equation first. If a variable became statistically insignificant to the model after the
addition of another variable, it was removed from the model. (Warner, 2013, p 560-561). Both
standardized (β) and unstandardized coefficients (B) were reported, but the discussion of the
models focused on the unstandardized coefficients as the predictor variables had already had
scores standardized by transformation into Z-scores. Squared semipartial correlations (sr²) were
reported and represented the variance of outcome scores that could be uniquely predicted by each
variable. For a predictor, sr² could be used to interpret effect size when all other predictors were
statistically controlled. The semipartial correlation value (sr) was compared with Cohens’s d
effect size guidelines (Cohen, 1988).

Large scale studies

The large scale studies included students of a range of prior knowledge levels.
Independent samples t-tests show that for all groups that completed the BRI and ROXCI, pre-test
scores were not statistically significant between the experimental and control groups. Statistical
regression analyses were performed for the results from the four large scale studies at a
confidence level of 90%. Two of the four large scale studies tested bonding representation
concepts (BRI), and two tested redox concepts (ROXCI). Variables tested to be predictors of
post-test scores included normalized pre-test scores, reading comprehension scores, and
participation in the intervention of reading a text passage about the topic (dummy scored, read
related text = 1, control group = 0). All interactions were also analyzed through regression for
each study. The summary of results for the four large scale studies is presented in Table 1.9. For
each study, the following linear regression model was tested:
\[ Z - score \ post - test = b_0 + b_1 (read \ related \ text) + b_2 (GMRT \ Z - score) + b_3 (pre - test Z - score) + b_4 (read \ text \star GMRT \ Z - score) + b_5 (read \ text \star pre - test Z - score) + b_6 (GMRT \ Z - score \star pre - test Z - score) + b_7 (read \ text \star GMRT \ Z - score \star pre - test Z - score) \]

**Study A**

Study A included a sample of 290 participants in the second semester of a traditional two-semester general chemistry sequence, and they completed the ROXCI as their pre- and post-test measures. The interaction between reading comprehension score and pre-test score was found to be statistically significant in the linear regression model (\( \alpha = 0.10 \)) (\( p = 0.054, sr^2 = 0.010 \)), and the coefficient was a positive value (\( B = 0.017 \)). The final statistical regression model (adjusted \( R^2 = 0.195 \), \( F(3,286) = 10.981, p < 0.001 \)) could explain approximately 20% of the variance of post-test scores, and it included the main effects of GMRT Z-score (\( p = 0.040, sr^2 = 0.012 \)) and pre-test Z-score (\( p < 0.001, sr^2 = 0.083 \)) in addition to the interaction discussed above. 1% and 8% of the variation of scores could be uniquely predicted by the main effects of reading skill and prior knowledge, respectively, and the effects are considered to have small to medium effect sizes. In this population of students, those who were more skilled readers were predicted to score higher than average readers on the post-test, statistically controlling for other variables, and higher prior knowledge students were also predicted to perform better than average students. The final regression model is expressed by the equation:

\[ Z - score \ on \ ROXCI \ post - test = -0.02 + 0.18 (GMRT \ Z - score) + 0.45 (pre - test Z - score) + 0.17 (pre - test Z - score \star GMRT \ Z - score) \]
Study B

The second large scale study was conducted with participants of a one-semester general chemistry course designed for engineering majors. This population (N = 143) completed the BRI. The results of the statistical linear regression differed from both studies described above. The final model (adjusted $R^2 = 0.275$, $F(2,141) = 8.641, p < 0.001$) could explain 28% of the variance of BRI post-test scores and included pre-test scores and the interaction between reading text and pre-test scores as the two statistically significant predictors. On average, students scored a $Z$-score of 0 on the post-test (by definition of a $Z$-score), but those with higher prior knowledge were predicted to score positive $Z$-scores ($p < 0.001, sr^2 = 0.129$). Students scoring one standard deviation above the mean on the pre-test were predicted to score 0.6 standard deviations above the mean on the post-test when controlling for other variables statistically.

The interaction between prior knowledge and reading the text was found to be statistically significant at 90% confidence ($p = 0.076, sr^2 = 0.016$). The negative correlation between this interaction and post-test scores suggests that for those who did read the related text, gains would be made for students with lower prior knowledge (pre-test $Z$-scores which are negative, or lower than the mean average), and losses would be made for students with higher prior knowledge (pre-test $Z$-scores above the mean average). This interaction has a low effect size based on the squared semipartial, and could uniquely predict 2% of the variance of post-test scores. The final model was:

$$Z - score\ on\ BRI\ post - test = -0.12 + 0.57(pre - test\ Z - score) - 0.29 (pre - test\ Z - score \ast read\ related\ text)$$
For example, if a student scored one standard deviation above the mean on the pre-test and did read the bonding text as learning intervention, the predicted post-test Z-score would be

\[ 0 + (0.57)(1) – (0.29)(1 * 1) = 0.28 \] standard deviations above the mean.

A student who also scored one standard deviation on the pre-test but did not read the related text would be predicted to have a post-test Z-score

\[ 0 + (0.57)(1) – (0.29)(1 * 0) = 0.57 \] standard deviations above the mean.

Based on this model, students with higher prior knowledge would not have learning gains upon reading the text, and would actually be predicted to score below those who did not read an expository text about the chemistry topic. However, students with lower prior knowledge would be helped by reading the related text. A student scoring one standard deviation below the mean on the pre-test (and who did not read related text) would have a predicted post-test score

\[ 0 + (0.57)(-1) – (0.29)(-1 * 0) = 0.57 \] standard deviations below the mean.

By reading related text, however, this student would be predicted to score 0.29 standard deviations closer to the mean average on the post-test (a final post-test Z-score of -0.28).

*Study C*

The population for Study C consisted of 396 students in the first-semester course of a traditional two-semester general chemistry sequence, and they were tested on bonding representations. The linear regression model that emerged from the analyses included all three main effects as significant predictors, and no interactions were significant at the \( \alpha = 0.10 \) level (adjusted \( R^2 = 0.188, F(3,395) = 12.826, p < 0.001 \)). Approximately 17% of the variance of BRI
post-test scores could be explained by the model. All three main effects had positive correlations with the BRI post-test. Students who read the text were predicted to score 0.30 standard deviations of the mean higher on the post-test than students who did not read the related text when reading ability and prior knowledge were statistically controlled (p = 0.001, $sr^2 = 0.023$). The effect size of this predictor was small.

Reading comprehension skill was found to statistically significantly predict post-test outcomes, but the effect size was small based on the squared semipartial ($sr^2 = 0.010$). Participants who were one standard deviation above the rest of the population in reading skill would be predicted to score 0.2 standard deviations above the class mean on the BRI post-test, controlling for prior knowledge and reading the intervention text (p = 0.033).

Finally, prior knowledge was found to have the largest effect size on post-test scores, with students one standard deviation above the mean on the pre-test predicted to score 0.41 standard deviations above the mean on the post-test (p < 0.001, $sr^2 = 0.055$). The linear regression analysis produced the final regression model:

$$Z - \text{score on BRI post-test} = -0.18 + 0.30(\text{whether related text was read}) + 0.16(\text{GMRT Z-score}) + 0.41(\text{pre-test Z-score})$$

**Study D**

The final large scale study was conducted with a one-semester general chemistry course designed for engineering students (N = 145), and the concept inventory used was the ROXCI. The final model included two statistically significant predictors at $\alpha = 0.10$ (whether related text was read and the interaction between reading text and prior knowledge), though the overall model was not statistically significant (adjusted $R^2 = 0.012$, F(2,144) = 1.258, p = 0.276).
1% of the variance in the ROXCI post-test could be predicted by the variables available in this study. The variables which were statistically significant, however, will still be discussed.

The text intervention was statistically significant \((p = 0.080, sr^2 = 0.021)\) with a small effect size based on the squared semipartial. The effect of reading the text, however, was opposite of the effect found in study C, as reading in this population had a negative correlation with the ROXCI post-test \((B = -0.24)\). Students who read the text, on average, scored 0.24 standard deviations below the mean (controlling for other variables), whereas students who did not read the text were predicted to score the mean average on the post-test.

However, when prior knowledge as measured by the pre-test was accounted for, an interaction was found between reading and prior knowledge \((p = 0.047, sr^2 = 0.028)\). The effect was similar to the effect of the interaction found in Study B, as there was a negative coefficient for the interaction. This interaction predicts that for someone with high prior knowledge (scored one standard deviation above the mean on the pre-test), reading the text about redox concepts will result in a predicted post-test score 0.37 standard deviations lower than the mean. Interestingly, the main effect of prior knowledge based on pre-test scores was not found to be predictive of post-test scores \((p = 0.391)\), even though this predictor was significant in the other three large scale studies. The final regression model was:

\[
Z - \text{score on BRI post - test} = -0.02 - 0.24(\text{whether text was read}) - 0.29(\text{pre - test Z - score} \ast \text{read text})
\]
Table 1.9 Statistical linear regression results for studies A, B, C, and D, with only statistically significant predictors included.

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>Standard error</th>
<th>β</th>
<th>t</th>
<th>sr²</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Study A: Chem 404, ROXCI (N=290)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.02</td>
<td>0.09</td>
<td>-0.22</td>
<td>0.829</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GMRT Z-score</td>
<td>0.18</td>
<td>0.09</td>
<td>0.18</td>
<td>2.07</td>
<td>0.012</td>
<td>0.040</td>
</tr>
<tr>
<td>Pre-test Z-score</td>
<td>0.45</td>
<td>0.08</td>
<td>0.45</td>
<td>5.46</td>
<td>0.083</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Interaction between GMRT and pre-test</td>
<td>0.17</td>
<td>0.09</td>
<td>0.17</td>
<td>1.93</td>
<td>0.010</td>
<td>0.054</td>
</tr>
<tr>
<td>$R^2 = 0.214$, Adjusted $R^2 = 0.195$, $R = 0.463$, $F(3, 286) = 10.981$, $p &lt; 0.001$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| **Study B: Chem 405, BRI (N=142)**            |       |                |       |       |       |       |
| Constant                                      | -0.12 | 0.11           | -1.06 | 0.292 |       |       |
| Pre-test Z-score                              | 0.57  | 0.11           | 0.56  | 5.01  | 0.129 | <0.001|
| Interaction between reading text and pre-test | -0.29 | 0.16           | -0.22 | -1.79 | 0.016 | 0.076 |
| $R^2 = 0.311$, Adjusted $R^2 = 0.275$, $R = 0.558$, $F(2, 141) = 8.641$, $p < 0.001$ |       |                |       |       |       |       |

| **Study C: Chem 403, BRI (N=396)**            |       |                |       |       |       |       |
| Constant                                      | -0.18 | 0.07           | -2.71 | 0.007 |       |       |
| Whether text was read                         | 0.30  | 0.09           | 0.15  | 3.31  | 0.023 | 0.001 |
| GMRT Z-score                                  | 0.16  | 0.07           | 0.14  | 2.14  | 0.010 | 0.033 |
| Pre-test Z-score                              | 0.41  | 0.08           | 0.37  | 5.11  | 0.055 | <0.001|
| $R^2 = 0.188$, Adjusted $R^2 = 0.173$, $R = 0.433$, $F(3, 395) = 12.826$, $p < 0.001$ |       |                |       |       |       |       |

| **Study D: Chem 405, ROXCI (N=145)**          |       |                |       |       |       |       |
| Constant                                      | -0.02 | 0.09           | -0.25 | 0.806 |       |       |
| Whether text was read                         | -0.24 | 0.13           | -0.15 | -1.76 | 0.021 | 0.080 |
| Interaction between reading text and pre-test | -0.37 | 0.18           | -0.24 | -2.01 | 0.028 | 0.047 |
| $R^2 = 0.060$, Adjusted $R^2 = 0.012$, $R = 0.246$, $F(2, 144) = 1.258$, $p = 0.276$ |       |                |       |       |       |       |
Effect of reading comprehension when text was read

If reading comprehension skill level affected the information students were able to glean from reading a text passage, an interaction term between reading the text and reading skill level would be present in the final linear regression equation. This would suggest that students who did read related text (read related text dummy variable = 1) would have their post-test scores on the inventories affected by their reading skill level differently than would students who did not read the text (did not read related text dummy variable = 0). Not one of the four large scale studies contained a statistically significant interaction between reading text and reading comprehension skill, suggesting that there is no additional affect from reading comprehension on post-test scores when a passage about the concept tested was read.

Meta-analysis

Description

Because the results of the four large scale studies were not consistent a meta-analysis was conducted to pool the results of Studies A-D.

When large-scale studies are used to test effects of interventions, the results of the studies are often untested with additional studies. If the results are investigated with a new population, there is the possibility that the outcomes and effects are different, and the researchers then face the difficult task of reconciling the disparate studies. Meta-analysis is a statistical technique for comparing and combining results from a number of studies. By analyzing the effect sizes of predictors in different studies, a weighted average effect size can emerge to produce a more generalizable effect size that may not have been apparent in any one study (Glass, 1976; Fitz-
Gibbon, 1986; Rosenthal & DiMatteo, 2001; Cumming, 2012). However, meta-analyses may be hindered by publication bias, as studies that do not present marked results tend not to be published (Rosenthal & DiMatteo, 2001, Borenstein et al., 2009, p 378-384). A small-scale meta-analysis can be conducted using results from a number of studies carried out by a research team. Using unpublished results will eliminate the danger of publication bias. Another criticism of meta-analysis is the heterogeneity of studies investigated. Including studies conducted by the same research group and with the same methodology will eliminate the concerns raised by heterogeneity of methodologies and analysis techniques.

A meta-analysis was performed to compare results from the four large-scale studies. The raw data in this analysis consists of the regression results from the four independent studies. In particular the effect sizes of each predictor (sr) and the sample size of the population. The effects were entered as fixed effects into the Exploratory Software for Confidence Intervals (ESCI) by Cumming & Calin-Jageman. For each predictor, the output of the analysis includes lower level (LL) and upper level (UL) confidence intervals at 95% confidence for the effect size from each study, the weight percentage of each study, and the overall weighted effect of the predictor (and the confidence intervals) (Cumming, 2012).

Results

The four effects that were found in a majority of the studies were included in the meta-analysis, including whether text was read, reading skill level (GMRT scores), prior knowledge (pre-test scores), and the interaction between reading a related text and prior knowledge). The results of the meta-analyses are presented in Table 1.10, and a Forest Plot of the results is presented in Figure 1.1. If a confidence interval overlaps with Pearson’s $r = 0$, then that effect is
interpreted as being non-statistically significant. The effect size was estimated using Cohen’s guidelines.

Overall, reading the text appears to have no effect on the post-test outcome. The weighted effect had a Pearson’s r value of 0.05, with a 95% confidence interval of -0.012 to 0.0114). This finding suggests that simply providing reading material about a concept will have no significant effect on learning gains when not accounting for other student characteristics.

The fixed effects model did find reading skill to have a statistically significant (but small) effect on post-test scores, with a semipartial r of 0.08 (CI = 0.021 – 0.147). Students with higher reading comprehension skill were predicted to have positive gains on their post-tests, controlling for other variables.

Prior knowledge had the largest effect size (a medium effect size), with a weighted semipartial r value of 0.246 (CI = 0.183 – 0.309). It is no surprise that pre-test score has a positive correlation with post-test score, statistically controlling for other variables. Students who have correct understanding of a concept during a pre-test will likely maintain that understanding during a post-test.

The final effect tested was the interaction between reading the concept text and prior knowledge. This effect was not found to be statistically significant at 95% confidence (semipartial r = -0.06, CI = -0.120 – 0.006) from the meta-analysis, but the results do have implications that will be addressed in the discussion.
Table 1.10 Data included in the meta-analysis and the Fixed Effects Model output results for the effects of whether text was read, reading skill level, prior knowledge, and the interaction between prior knowledge and reading text.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Pearson’s r (sr)</th>
<th>LL (95% confidence)</th>
<th>UL (95% confidence)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Whether text was read</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A (N=290)</td>
<td>0.18</td>
<td>-0.115</td>
<td>0.115</td>
</tr>
<tr>
<td>B (N=142)</td>
<td>0.07</td>
<td>-0.096</td>
<td>0.162</td>
</tr>
<tr>
<td>C (N=396)</td>
<td>0.15</td>
<td>0.054</td>
<td>0.247</td>
</tr>
<tr>
<td>D (N=145)</td>
<td>-0.15</td>
<td>-0.302</td>
<td>0.017</td>
</tr>
<tr>
<td>Fixed Effects Model</td>
<td>0.05</td>
<td>-0.012</td>
<td>0.114</td>
</tr>
<tr>
<td><strong>Reading skill</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0.11</td>
<td>-0.006</td>
<td>0.221</td>
</tr>
<tr>
<td>B</td>
<td>-0.01</td>
<td>-0.177</td>
<td>0.152</td>
</tr>
<tr>
<td>C</td>
<td>0.10</td>
<td>-0.001</td>
<td>0.195</td>
</tr>
<tr>
<td>D</td>
<td>0.90</td>
<td>-0.075</td>
<td>0.248</td>
</tr>
<tr>
<td>Fixed Effects Model</td>
<td>0.09</td>
<td>0.021</td>
<td>0.147</td>
</tr>
<tr>
<td><strong>Prior knowledge</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0.29</td>
<td>0.179</td>
<td>0.390</td>
</tr>
<tr>
<td>B</td>
<td>0.36</td>
<td>0.206</td>
<td>0.494</td>
</tr>
<tr>
<td>C</td>
<td>0.23</td>
<td>0.139</td>
<td>0.325</td>
</tr>
<tr>
<td>D</td>
<td>0.07</td>
<td>-0.093</td>
<td>0.231</td>
</tr>
<tr>
<td>Fixed Effects Model</td>
<td>0.246</td>
<td>0.183</td>
<td>0.309</td>
</tr>
<tr>
<td><strong>Interaction between reading text and prior knowledge</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>-0.05</td>
<td>-0.165</td>
<td>0.065</td>
</tr>
<tr>
<td>B</td>
<td>-0.13</td>
<td>-0.287</td>
<td>0.038</td>
</tr>
<tr>
<td>C</td>
<td>0.00</td>
<td>-0.095</td>
<td>0.103</td>
</tr>
<tr>
<td>D</td>
<td>-0.17</td>
<td>-0.320</td>
<td>-0.003</td>
</tr>
<tr>
<td>Fixed Effects Model</td>
<td>-0.06</td>
<td>-0.120</td>
<td>0.006</td>
</tr>
</tbody>
</table>
Figure 1.1 Forest plots for the four predictors included in the meta-analysis. The weighted mean Pearson’s r for each effect has a bolded outline. The effects are (blue circles) whether text was read, (orange squares) reading comprehension skill, (grey triangles) measure of prior knowledge (pre-test scores), and (yellow diamonds) the interaction between reading text and prior knowledge.

**Controlled**

*Did reading text affect post-test scores?*

For the controlled experiment, the variances between the group that read the text about bonding representations concepts and the group that read the text about loons (the control group) were equal (Levene’s test for equality of variances $p = 0.834$), and the mean average pre-test scores were not statistically different at $\alpha = 0.05$ ($p = 0.586$). However, the mean post-test scores differed by 4.1 points ($p < 0.001$), with the group reading the bonding text averaging a higher post-test score (Table 1.11). This finding suggests that reading text may have a positive effect on
learning gains for this controlled population. Comparisons of pre- and post-test scores between the groups who did read the text and who did not read the text are shown in Figure 1.2.

Table 1.11 Independent samples t-test results (α = 0.05) comparing BRI post-test raw scores between those who read the bonding text and those who did not read the bonding text.

<table>
<thead>
<tr>
<th></th>
<th>Students who did read text</th>
<th>Students who did not read</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>38, 18</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>12.5 (8.3)</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>3.9 (3.1)</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>4.030</td>
<td></td>
</tr>
<tr>
<td>η²</td>
<td>0.23</td>
<td></td>
</tr>
</tbody>
</table>

What variables predicted BRI post-test Z-scores?

A stepwise statistical regression was conducted using SPSS to determine whether any variables emerged as statistically significant predictors of post-test Z-scores. The model tested the following equation:

\[
Z - score\ on\ BRI\ post - test = b_0 + b_1(\text{read text}) + b_2(\text{GMRT Z - score}) + b_3(\text{pre - test Z - score}) + b_4(\text{read text} \ast \text{GMRT Z - score}) + b_5(\text{read text} \ast \text{pre - test Z - score}) + b_6(\text{GMRT Z - score} \ast \text{pre - test Z - score}) + b_6(\text{read text} \ast \text{GMRT Z - score} \ast \text{pre - test Z - score})
\]
Figure 1.2 BRI pre-test scores (left) for students who did read text (green) and who did not read text (blue patterned). On the left are BRI post-test scores for students who did read text (green) and who did not read text (blue patterned).

The results showed that three predictors were considered statistically significant at the 95% confidence level: reading the concept text passage, the interaction between pre-test scores and reading related text, and the interaction between pre-test scores and reading comprehension scores (adjusted $R^2 = 0.445$, $F(3,52) = 8.721$, $p < 0.001$), and the results are presented in Table 1.12. The model from the analysis was able to explain approximately 45% of the variance of post-test scores. The mean average BRI post-test Z-score for the sample was -0.709 ($p < 0.001$), controlling for pre-test, reading ability, and whether the bonding text was read. Participants who read the text about bonding representations scored, on average, 1.11 standard deviations higher on the post-test than students who did not read the text ($p < 0.001$, $sr^2 = 0.268$). Approximately 27% of the variance of post-test scores could be uniquely predicted by whether the participant...
read text, controlling for all other variables, and reading had a large effect size, based on the squared semipartial.

As noted above, students in this sample scored, on average, 30 points out of a possible 48 points on the GMRT, but scores ranged from 6/48 to 47/48. The main effect of reading comprehension score, however, was not statistically significant in the final linear regression model (p = 0.832). In this sample, reading comprehension skill alone did not impact the post-test BRI score. However, the interaction between reading comprehension skill and pre-test score was a statistically significant indicator. Students who scored one standard deviation above the mean on the GMRT and one standard deviation above the mean were predicted to score 0.2 standard deviations lower on the post-test than students who scored mean average reading and pre-test scores (p = 0.035, $sr^2 = 0.047$). Based on the squared semipartial this interaction had a small effect size, and approximately 5% of the variance of post-test scores could be uniquely predicted by the interaction. The negative coefficient for this interaction has interesting implications. This could mean that students with lower than average prior knowledge (a negative Z-score on the BRI pre-test), but higher than average reading ability (a positive Z-score on the GMRT) would make modest learning gains on the post-test, controlling for whether text was read. However, students with higher than average prior knowledge but lower than average reading skill would be predicted to have the same gains.

The third statistically significant predictor of BRI post-test Z-scores was the interaction between prior knowledge and reading the bonding text. Students who read the text and scored one standard deviation above the mean on the BRI pre-test were predicted to score 0.536 more standard deviations above the mean than similar students who did not read the text (p < 0.001, $sr^2 = 0.171$). 17% of variance in post-test scores was uniquely explained by the interaction. The
interaction had a medium effect size, based on the squared semipartial. This finding suggests that students with more prior knowledge were helped more by reading the text than students with less prior knowledge students, regardless of reading comprehension ability. Notably, BRI pre-test score was not a statistically significant main effect, likely because the students in this controlled study all scored below the mean average on the pre-test, and so the range of pre-test scores in this sample was small. The overall linear regression equation was:

\[
Z - \text{score on BRI post-test} = -0.71 + 1.1(\text{read related text}) \\
- 0.20(\text{pre-test Z-score} \times \text{GMRT Z-score}) + 0.54 (\text{pre-test Z-score} \times \text{read related text})
\]

Table 1.12 Statistical stepwise regression analysis results for predictors of BRI post-test Z-scores for controlled group.

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>Standard error</th>
<th>β</th>
<th>t</th>
<th>sr²</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant (b₀)</td>
<td>-0.709</td>
<td>0.176</td>
<td>-4.038</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whether related text was read (1=\text{did read}, 0=\text{did not read})</td>
<td>1.11</td>
<td>0.215</td>
<td>0.522</td>
<td>5.161</td>
<td>0.268</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Interaction between reading and pre-test score</td>
<td>0.536</td>
<td>0.130</td>
<td>0.417</td>
<td>4.120</td>
<td>0.171</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Interaction between pre-test score and reading comprehension score</td>
<td>-0.201</td>
<td>0.093</td>
<td>-0.221</td>
<td>-2.164</td>
<td>0.047</td>
<td>0.035</td>
</tr>
</tbody>
</table>

\(R^2 = 0.476, \text{Adjusted } R^2 = 0.445, R = 0.690, F(3, 52) = 8.721, p < 0.001\)
Discussion and conclusions

Individual studies

When the five studies are investigated separately, the resulting predictors of learning gains (measured by a pre-/post-test design) differ for each experiment. In fact, no two studies, when analyzed by statistical linear regression, had the same group of statistically significant predictors. However, in each study, the means of the pre- and post-test scores did not differ overall. There appears to be no testing effect present in these studies. Once subpopulations were compared, however, some predictors were statistically significant more often than others, including whether the text was read (Controlled, Study C, and Study D), prior knowledge (Study A, Study B, and Study C), reading comprehension skill (Study A and Study C) and an interaction between prior knowledge and reading the concept text (Controlled, Study B, and Study D).

Prior knowledge and reading comprehension skill have been found to be predictive of chemistry course performance (Pyburn & Pazicni, 2014, Akbaşlı et al., 2016). The interaction between reading text and prior knowledge, however, has interesting implications on pedagogical strategies. As illustrated by the final regression models, students who read the concept text did not all make the same gains in learning as assessed by the BRI or ROXCI. Students who scored below the mean average on the pre-test (low prior knowledge) did not catch up to students who scored above the mean on the pre-test (high prior knowledge), but they were predicted to close the knowledge gap. Of particular note is that this interaction was seen when different chemistry concepts were tested. The effect is not specific to just one concept inventory or one text passage, but results appear to be generalizable to multiple concepts.

Both populations who displayed the interaction were in the general chemistry course designed for engineers. These students, on average, have more science background coming in to
general chemistry at the university level than do the students in the traditional general chemistry course. This interaction was not observed at the 95% confidence level for the traditional general chemistry courses, but the effect may be hidden by the large amount of variance of scores in those populations.

Even if a variable is found to not be statistically significant at $\alpha = 0.05$, the variable may still have an effect on the outcome measure that is undetected by the statistical analyses chosen for the experiment. On the other hand, even if a variable is found to be significant at that confidence level, the effect that predictor has on the outcome may be unimportant, particularly if the sample size is large (Olejnik & Algina, 2000; Glass & Cohen, 2012; Cumming, 2012). In order to better understand the results of the linear regression analyses, effect sizes must be compared through meta-analysis.

**Meta-analysis**

*Prior knowledge and confidence intervals*

The variable with the largest effect size for predicting concept post-test scores was prior knowledge. This result comes as no surprise as linear regression analyses of three of the four large scale studies found this main effect to be statistically significant with medium effect sizes. The meta-analysis performed using ESCI uses the degree of overlap between confident intervals of Pearson’s r values for a predictor (Cumming 2012). 95% confidence intervals were computed for the correlation of prior knowledge scores with the post-test scores of each large study. As shown in Figure 1.1, the confidence intervals overlap for all four studies. The confidence interval for Study A can be interpreted in the following way: If 100 additional samples were taken for Study A, the true population mean for the correlation between prior knowledge and post-test score would fall between 0.179 and 0.390 95 times. When there is significant overlap between
confidence intervals of several studies, the power of the calculated mean correlation becomes larger. That is, the probability of making a Type II error decreases. Even though Study D did not find prior knowledge to be statistically significant, the confidence interval for Pearson’s $r$ was wide enough to overlap with the confidence intervals of the three other studies, thus enhancing the power of the $r$ statistic.

*Reading comprehension*

Reading comprehension scores were found to be statistically significant positive predictors of post-test outcome, with small and medium effect sizes, respectively. Again, there was sufficient overlap among all four confidence intervals of the Pearson’s $r$ correlation between reading skill and post-test scores that the power of the statistic was enhanced by comparing the effect from each study. So although the effect of reading skill was not statistically significant in half the studies, the confidence intervals calculated at 95% confidence overlapped to a degree where the probability of not finding an effect when there is an effect will decrease. The interpretation from this study is that, although a correlation between reading skill and post-test score was found to not be significant in individual studies, the probability of making a Type II error may be large, and when the results of the studies are pooled, the effect emerges as significant. Pyburn *et al.* found that reading comprehension skill correlated positively with course performance (2013), and the results of the meta-analysis of the four large-scale studies support this finding. Prior knowledge had a medium effect size, the largest of the variables analyzed (semipartial $r = 0.246$), but reading comprehension skill may help compensate to a small degree for low prior knowledge (semipartial $r = 0.086$). A model of comprehension ability put forward by Gernsbacher *et al.* (1990) posits that students who are more adept at
comprehending text can more easily make connections between ideas while suppressing errant information, and thus can begin to structure understanding of concepts more efficiently. Readers with lower comprehension skill can get bogged down by the language used in the assessments or text, and will then require more effort to build connections between topics and concepts.

*Interaction between prior knowledge and reading*

The interaction between pre-test score and reading the text, though not statistically significant at 95% confidence, begins to emerge as a possible predictor in the meta-analysis. The confidence interval only passes the $r = 0$ mark to $r = 0.006$, a very slight overlap. With more measurements this effect may become statistically significant at 95% confidence. The implication of such an interaction warrants some discussion.

The meta-analysis of the four large-scale studies conducted suggests reading a text passage may differentially affect students based on prior knowledge, as evidenced by the interaction of prior knowledge and reading the concept text. Students who perform lower than average on the pre-test and then read the text passage about the chemistry topic will have a small gain in their post-test score, whereas higher performing students who read the text passage will have a small reduction in their post-test score, controlling for all other variables. Though the interaction does not have a large enough effect ($r = -0.06$) for reading related text to completely compensate for low prior knowledge, the implication is that reading a text passage about a concept can aid learning for students with little prior knowledge about the concept.

This effect (termed Expertise Reversal) has been observed in previous studies (Kalyuga, 2007; Kalyuga & Renkl, 2010), in which teaching techniques result in different outcomes based on the level of prior knowledge of the learner. Cognitive load theory describes how working
memory load can affect a learner’s ability to process and retain information (Sweller, 1988). Text passages contain new vocabulary, examples, definitions, and arguments that the reader must navigate in order to process the main concepts presented. Highly cohesive and redundant materials are more useful for novice learners, but those with high prior knowledge may experience a higher cognitive load when encountering extraneous information. Those students with higher prior knowledge, in order to decrease the cognitive load, may then begin to merely skim the reading material, and can miss important pieces of information. Students with low prior knowledge will become more engaged with the cohesive texts, and may then be able to retain more nuanced information from the cohesive text. Thus, the characteristics of a text passage need to be designed with the level of prior knowledge of students in mind.

The text passages in this study were written such that the readability dimensions matched the average characteristics of passages found in popular chemistry textbooks (Pyburn & Pazicni, 2014). However, there was a very high level of referential cohesion in both text passages relative to traditional chemistry textbook passages, and this level of cohesion may have benefitted those with low prior knowledge more than those with high prior knowledge, leading to the small expertise reversal effect which begins to emerge from the data upon meta-analysis.

*Revisiting Research Question 1: To what degree will reading about a general chemistry concept promote learning gains?*

The results of the meta-analysis show that reading on its own does not significantly increase post-test scores on concept inventories. However, reading text is moderated by individual characteristics – namely, prior knowledge. Prior knowledge on its own has the
greatest correlation with post-test performance \((r = 0.246)\), with a medium effect size. However, students with lower prior knowledge are able to make small gains in order to lessen the knowledge gap when they read a highly cohesive text. Students with high prior knowledge, however, may experience a higher cognitive load because of redundancies in the text, and will therefore be less engaged with the text in order to lessen the cognitive load, and will have the tendency to miss new information upon reading.

*Revisiting Research Question 2: How does student reading comprehension influence learning gains when a text passage is used for studying a concept?*

Based on the results of the individual studies and the meta-analysis, reading skill has a small effect on learning gains, but not necessarily because text about the concept was read. No statistically significant interaction between reading the text and reading comprehension skill emerged from the individual studies or from a meta-analysis of the effect sizes of that interaction. The main effect of reading skill affected both readers and non-readers alike in that there is a small correlation between reading skill and post-test scores \((sr = 0.09, LL = .021, UL = 0.147)\). Previous studies have shown a positive relationship between science achievement and reading skill (O’Reilly & McNamara, 2007; Pyburn *et al.*, 2013; Reed *et al.*, 2016). This study contributes to those conclusions, showing a positive, though modest, effect of reading comprehension ability on performance on chemistry concept inventories.
Revisiting Research Question 3: Does an interaction between reading comprehension and prior knowledge predict learning outcomes, and to what extent?

No interaction between reading comprehension skill and prior knowledge exists in the populations sampled for these studies. An Expertise Reversal Effect, however, may have a small effect on the post-test outcome. This is an interaction between prior knowledge and reading a related text. This effect, though not statistically significant at 95% confidence, may become more pronounced if different types of texts are studied. Having texts with a range of cohesion scores would allow for the effect to affect more of the population. In this study, because the text cohesion was very high, an expertise reversal would only be noticeable for low prior knowledge students to have gains. To really examine whether this effect occurs in this population, a text that may differentially aid high prior knowledge students (one with low cohesion scores) should be included in any future studies.

Implications

Expertise Reversal and text cohesion

The implications of an expertise reversal effect when reading text in order to learn about a concept are that individuals must be aware of their own characteristics to choose the best and most effective materials for studying. If only the same study materials are provided to all students, both high prior knowledge and low prior knowledge would suffer. Having a variety of texts which differ in text cohesion would be one way to provide the optimal learning opportunities for a wide range of students.

It must be noted, however, that in order for a student to choose the best material for him or herself, they must be aware of their level of knowledge about the concept. Providing an
opportunity for students to take a pre-test or to complete a concept inventory in order to assess their level of prior knowledge would be absolutely necessary.

**Meta-analysis as a research tool**

The power of meta-analysis to increase the statistical power of effects has been used for very large scale studies including tens or hundreds of literature sources. However, the technique is not limited to such a large scale. In this study only four experiments were compared to find statistically significant effects that would have been lost in the noise of variance without the meta-analysis. Though a researcher may be able to spot a trend in the data, this trend does not always pan out to be statistically significant when analyzed through traditional methods. By effectively increasing the overall sample size, previously non-significant effects will increase in statistical power. This can be a particularly powerful tool when variables have small effect sizes and when several studies have disparate results.

Also, the method of using meta-analysis emphasizes the importance of effect size over p values when evaluating the significance of effects. In very large samples, predictors can be considered statistically significant even when contributing very little to the magnitude of the outcome. However, effect sizes measured by Pearson’s r or Cohen’s d display more clearly the magnitude of difference of means or the magnitude of the phenomenon.

**Limitations of study**

A large limitation of this study was the very high level of cohesion found in both concept texts. This text characteristic may have enhanced an expertise reversal effect that may not have otherwise been present. However, the high cohesion may have highlighted the expertise reversal effect which may not have otherwise been detected. Future studies should include different text
types to further investigate whether text cohesion differently affects students based on level of prior knowledge.

Another limitation is comparing the Gates-MacGinitie Reading Test results to learning outcomes based on reading. The text passages used in this study have very low narrativity, and are more expository in nature. However, more passages in the 7/9th grade GMRT test were found to be narrative in nature than expository based on a study by Rowe et al. (2006). It is likely that the 10/12th grade version of the GMRT follows the same trend, and so using a reading comprehension measure that focuses more on narrative text to analyze the relationship between reading skill and learning using an expository text may limit the interpretation of results.

Implementation of the findings of this study in a classroom may be difficult. Having the appropriate reading materials available for students of differing levels of prior knowledge for all concepts taught during general chemistry would be a daunting task for authors of the reading material.
CHAPTER 2

INVESTIGATION OF EPISTEMOLOGICAL BELIEFS OF GENERAL CHEMISTRY STUDENTS

Objectives

Chapter 1 discusses how text characteristics may differentially aid students depending on student prior knowledge. Another student trait that may affect why students may turn to textbooks, or use particular features to study, is their epistemological beliefs (EB). General chemistry students often view themselves as passive learners whose responsibility is to remember and memorize what their lecturers or textbooks say about a topic (Mazzarone & Grove, 2013). This understanding of the learner as a passive participant in learning, where knowledge is something an expert must hand down, demonstrates a naïve level of epistemological beliefs. This level of understanding about the source of knowledge is labeled as “Absolute Knowing” by Baxter Magolda (2004). However, students who recognize that they have some responsibility and ownership of their own learning would hold more sophisticated levels of epistemological beliefs. A more sophisticated level of belief is “Transitional Knowing,” where the individual believes that some areas of knowledge may be uncertain while other areas are concrete. Those who hold these beliefs may still recognize the importance of an expert for imparting knowledge, but are beginning to understand that some knowledge can be obtained by the individual him/herself. The next step in level of sophistication is “Independent Knowing.” Those who hold this level of epistemological understanding believe that most knowledge is uncertain and that personal experience is more important for understanding than a “correct” answer. The final level of understanding, according to Baxter Magolda (2004) is that of
“Contextual Knowledge,” where all knowledge exists in context, and is therefore uncertain. Evidence is crucial for the development of knowledge, and that an individual is responsible for their own understanding of that knowledge. The goal of this study was to use a quantitative epistemological beliefs survey to identify the levels of EB present among a population of students enrolled in the first semester of a two semester general chemistry course sequence. Specifically, the following questions were explored:

1. How sophisticated are epistemological beliefs about physical science held by first-semester general chemistry students, and are there differences between male and female students?

Previous studies have found that people may hold different levels of EB for different domains. A physical science centered EP assessment (Epistemological Beliefs Assessment for Physical Sciences, EBAPS) focuses on dimensions of EB that are of particular importance in science.

2. To what degree do levels of EB correlate with course performance, sex, and instructor?

Is student understanding of foundational chemistry topics affected by the student’s understanding of the source or structure of scientific knowledge? Course performance can be compared to epistemological beliefs (in terms of score on the EBAPS) to identify a relationship. The sex of the student and the lecturer of the course may also influence EB. These relationships will be investigated.

3. How do epistemological beliefs change over the course of a student’s first semester taking chemistry?
Can the short timeframe of one semester be enough time for a student to start to experience a change in their epistemological beliefs?

**Methodology**

**Setting**

Participants in this study were students enrolled in the first semester of a two-semester general chemistry sequence at a 4-year public research university with high research output in the Northeastern United States. Students were enrolled in one of three lecture sections. Two sections were taught by Lecturer A, and these sections met three times a week for 50 minutes each session. The third section was taught by Lecturer B, and this section met twice a week for 80 minutes each section.

**Study design**

In the third week of the semester students were invited to take an online survey about their beliefs about science knowledge (the EBAPS). Participation in the study was optional. Students completed the survey using the online data analysis program Qualtrics, and response time took an average of 15 minutes. During the final week of classes for the semester, students were once again invited to complete the same survey. Survey responses were coded and scored according to the scoring guide for the EBAPS. The overall scores and subscale scores were then compared with course performance scores using linear regression analysis and ANOVA with SPSS. Because no difference in EBAPS scores between lecture sections are significant, multilevel or hierarchical models were unnecessary for these data.
Outcome measures

Epistemological beliefs

The Epistemological Beliefs Assessment for Physical Sciences (EBAPS) is a 30-item forced-choice instrument developed by Elby and coworkers at the University of California, Berkeley (2006). Each item in the assessment is scored on a scale of 0-5. The scale is non-linear to account for variances of level of sophistication for each answer. The maximum score possible for the EBAPS is 120 points.

The assessment includes five subscales to probe student views among five non-orthogonal dimensions. The first subscale is “Structure of scientific knowledge” (ten items), which measures beliefs about whether knowledge consists of weakly connected facts and formulas or whether it is a highly-structured, coherent whole. The second subscale is “Nature of knowing and learning” (eight items). This subscale probes beliefs about whether learning consists of absorbing information or constructing one’s own understanding by practicing the material and reflecting on experiences. A third subscale is “Real-life applicability” (four items). This subscale teases out views of the applicability of science concepts as distinct from the desire to apply science in real life. The fourth subscale, “Evolving knowledge” (three items), examines whether one believes all scientific knowledge is set in stone (absolutism), there are no distinctions between evidence-based reasoning and mere opinion (extreme relativism), or some belief that spans the two extremes. The final subscale, “Source of ability to learn” (five items), measures understanding of whether someone is naturally good at science or whether most people are able to learn through hard work and doing science. Some items on the EBAPS belong to more than one subscale, as the item may touch on a number of beliefs.
There are three sections to the EBAPS. The first includes 17 items with statements that the student agrees/disagrees with on a five point Likert-scale. The second section is made up of six multiple choice items, and the third section is made up of seven items where the student indicates what side of a short debate he or she agrees with. The overall score on the EBAPS and the scores on the subscales were reported in terms of percentages. Students completed the EBAPS online during the third week and the final week of the semester.

Reliability of the EBAPS was not measured using a Crohnbach’s alpha statistic. The authors of the assessment note that the assessment items were designed so that students were allowed to disagree with themselves within a subscale. Because epistemological beliefs may be triggered depending on context, providing different contexts within the same subscales would allow the assessment to probe more nuanced, context-based beliefs of the students (Hammer & Elby, 2003).

**Course performance**

The measures of course performance were student scores on the first formal classroom assessment in the course and the final course grade at the end of the semester. As the participants were enrolled in three different course sections with two different lecturers, the course performance scores were transformed into Z-scores so that comparable values were available for all participants.

**Participants**

Participants’ demographic data (Table 2.1) were collected from institutional records. The population was predominantly female (84.4%), and 66.3% of the population were in their first
year. The most represented major was Biomedical Science: Medical and Veterinary Science (19.9%).

Table 2.1 Demographic statistics for participants.

|                           | % of students for whom these data are available | 84.4% Female  
Class standing       | 100%                                      | 15.6% Male  
Academic major        | 93%                                      | 66.3% First-years  
                        |                                          | 24.5% Sophomores  
                        |                                          | 3.4% Juniors  
                        |                                          | 1.9% Seniors  
                        |                                          | 19.9% Biomedical Science: Medical and Veterinary Sciences  
                        |                                          | 18.8% Biology  
                        |                                          | 5.7% Undeclared  
                        |                                          | 5.0% Biochemistry, Molecular and Cellular Biology  
                        |                                          | 4.6% Animal Sciences  
                        |                                          | 4.2% Medical Laboratory Sciences  
                        |                                          | 3.1% Chemistry

Data analysis and results

All univariate outliers were eliminated from analyses using a Mahalanobis distance criterion of $p < 0.001$ (Tabachnick & Fidell, 2013, p 99). Three outliers were omitted from the population of participants at the beginning of the semester ($N = 260$), and one outlier was omitted from the population at the end of the semester ($N = 120$).

Dependent and independent variables

Linear regression was used to investigate any relationships between epistemological belief scores (EBAPS total and subscores) and sex, lecturer, and course performance (the first midterm exam score and the score earned on the final, both normalized to Z-scores). Before the
regression analysis was performed, exploratory analysis of each variable was conducted. All continuous variables had skewness and kurtosis values between ±1, so all inferential statistics performed in this study were robust to the modest deviations from normality (Cohen et al., 1993, p 41).

Descriptive statistics for EBAPS scores

A total of 257 participants (excluding outliers) completed the EBAPS approximately three weeks into the fall semester of the first-semester course of a traditional 2-semester sequence general chemistry program. The mean average score was 80.0 points out of 120 possible points (66.7%), with a standard deviation of 10.7 points (8.9%). Descriptive statistics for total and subscale scores are presented in Table 2.2.

At the end of the same semester, 119 participants completed the EBAPS (80 total students completed the assessment at both time points). The mean score out of 120 points was 77.4 (64.5%), with a standard deviation of 11.5 (9.6%). Table 2.3 presents the descriptive statistics for EBAPS total and subscale scores for the end-of-semester assessment.

Subscale 1 scores at the beginning of the semester averaged 23.5 points out of 40 (58.8%), with a standard deviation of 4.4 points. At the end of the semester, students scored an average of 22.5 points (56.3%) on Subscale 1, with a standard deviation of 4.3 points. Subscale 2 scores at the beginning of the semester averaged 20.5 points out of 32 (63.8%), with a standard deviation of 4.0 points. At the end of the semester, students scored an average of 20.0 points (62.2%) on Subscale 2, with a standard deviation of 4.6 points. Subscale 3 scores at the beginning of the semester averaged 12.0 points out of 16 (73.8%), with a standard deviation of 2.6 points. At the end of the semester, students scored an average of 11.4 points (71.3%) on
Subscale 3, with a standard deviation of 2.7 points. Subscale 4 scores at the beginning of the semester averaged 8.3 points out of 12 (69.2%), with a standard deviation of 2.1 points. At the end of the semester, students scored an average of 7.9 points (65.8%) on Subscale 4, with a standard deviation of 2.4 points. Finally, Subscale 5 scores at the beginning of the semester averaged 15.6 points out of 20 (78.0%), with a standard deviation of 3.1 points. At the end of the semester, students scored an average of 15.3 points (76.5%) on Subscale 5, with a standard deviation of 3.1 points.
Table 2.2 Descriptive statistics of Epistemological Beliefs Assessment for Physical Science (EBAPS) scores at the beginning of the semester following removal of univariate outliers.

<table>
<thead>
<tr>
<th></th>
<th>Total Score</th>
<th>Subscale 1 (Structure of scientific knowledge)</th>
<th>Subscale 2 (Nature of knowing and learning)</th>
<th>Subscale 3 (Real-life applicability)</th>
<th>Subscale 4 (Evolving knowledge)</th>
<th>Subscale 5 (Source of ability to learn)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>257</td>
<td>257</td>
<td>257</td>
<td>257</td>
<td>257</td>
<td>257</td>
</tr>
<tr>
<td>Mean</td>
<td>80.0 (66.7%)</td>
<td>23.5 (58.8%)</td>
<td>20.4 (63.8%)</td>
<td>11.8 (73.8%)</td>
<td>8.3 (69.2%)</td>
<td>15.6 (78.0%)</td>
</tr>
<tr>
<td>Standard</td>
<td>10.7 (8.9%)</td>
<td>4.4 (11.0%)</td>
<td>4.0 (12.5%)</td>
<td>2.6 (16.2%)</td>
<td>2.1 (17.5%)</td>
<td>3.1 (15.5%)</td>
</tr>
<tr>
<td>deviation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.025</td>
<td>0.079</td>
<td>-0.106</td>
<td>-0.405</td>
<td>-0.127</td>
<td>-0.598</td>
</tr>
<tr>
<td>(Std. Error 0.152)</td>
<td>(Std. Error 0.152)</td>
<td>(Std. Error 0.152)</td>
<td>(Std. Error 0.152)</td>
<td>(Std. Error 0.152)</td>
<td>(Std. Error 0.152)</td>
<td>(Std. Error 0.152)</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-0.364</td>
<td>-0.437</td>
<td>0.086</td>
<td>-0.096</td>
<td>-0.305</td>
<td>-0.327</td>
</tr>
<tr>
<td>(Std. Error 0.303)</td>
<td>(Std. Error 0.303)</td>
<td>(Std. Error 0.303)</td>
<td>(Std. Error 0.303)</td>
<td>(Std. Error 0.303)</td>
<td>(Std. Error 0.303)</td>
<td>(Std. Error 0.303)</td>
</tr>
</tbody>
</table>
Table 2.3 Descriptive statistics of Epistemological Beliefs Assessment for Physical Science (EBAPS) scores at the end of the semester following removal of univariate outliers.

<table>
<thead>
<tr>
<th></th>
<th>Total Score</th>
<th>Subscale 1 (Structure of scientific knowledge)</th>
<th>Subscale 2 (Nature of knowing and learning)</th>
<th>Subscale 3 (Real-life applicability)</th>
<th>Subscale 4 (Evolving knowledge)</th>
<th>Subscale 5 (Source of ability to learn)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>119</td>
<td>119</td>
<td>119</td>
<td>119</td>
<td>119</td>
<td>119</td>
</tr>
<tr>
<td>Mean</td>
<td>77.4 (64.5%)</td>
<td>22.5 (56.3%)</td>
<td>19.9 (62.2%)</td>
<td>11.4 (71.3%)</td>
<td>7.9 (65.8%)</td>
<td>15.3 (76.5%)</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>11.5 (9.6%)</td>
<td>4.3 (10.8%)</td>
<td>4.6 (14.4%)</td>
<td>2.7 (16.9%)</td>
<td>2.4 (20.0%)</td>
<td>3.1 (15.5%)</td>
</tr>
<tr>
<td>Skewness</td>
<td>0.254</td>
<td>0.366</td>
<td>0.055</td>
<td>-0.122</td>
<td>-0.250</td>
<td>-0.231</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-0.728</td>
<td>-0.137</td>
<td>-0.804</td>
<td>-0.808</td>
<td>-0.304</td>
<td>-0.954</td>
</tr>
</tbody>
</table>

(Std. Error 0.222) (Std. Error 0.222) (Std. Error 0.222) (Std. Error 0.222) (Std. Error 0.222) (Std. Error 0.222)
Descriptive statistics for course performance

Measurements for course performance come from the first midterm exam and the final exam for each class section. Scores were transformed into Z-scores because two class sections completed different exams than the third class section with different scaling, and each assessment had different mean averages and standard deviations. In order to compare performance across the three subpopulations, a normalized Z-score was used as the performance score. The values are the number of standard deviations from the mean for each particular class section’s exam score.

The overall mean Z-score for the first midterm for all participants (N=294) was 0.097 with a standard deviation of 0.94. This mean average shows that the students who volunteered to participate in the study had scores that well represented the class as a whole because the Z-score is near zero and the standard deviation is near one. By definition, a mean average Z-score for a population is zero, with a standard deviation of one. The mean average Z-score for the final exam (N=296) was 0.14 (±0.92). The participants in this study on average scored 0.14 standard deviations above the class means of their final exam. A summary of the descriptive statistics for course performance measurements is presented in Table 2.4.

Table 2.4 Descriptive statistics of midterm and final grades (normalized as Z-scores).

<table>
<thead>
<tr>
<th></th>
<th>First midterm</th>
<th>Final</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N</strong></td>
<td>294</td>
<td>296</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>0.097</td>
<td>0.143</td>
</tr>
<tr>
<td><strong>Standard deviation</strong></td>
<td>0.941</td>
<td>0.924</td>
</tr>
<tr>
<td><strong>Skewness</strong></td>
<td>-0.493 (Std. Error 0.142)</td>
<td>-0.011 (Std. Error 0.142)</td>
</tr>
<tr>
<td><strong>Kurtosis</strong></td>
<td>-0.187 (Std. Error 0.283)</td>
<td>-0.011 (Std. Error 0.282)</td>
</tr>
</tbody>
</table>
Exploratory Analysis

Change in EBAPS scores

As stated above, three research questions guided this study. The first addresses the sophistication level of epistemological beliefs held by general chemistry students. The overall EBAPS score and the scores for the subscales were analyzed in order to interpret epistemological beliefs presented by Baxter Magolda’s model of epistemological reflection (2004). The scoring scheme for the EBAPS is gives answers to items point values between zero and five, with lower scores representing more naïve beliefs (Absolute Knowing or Transitional Knowing) and higher scores representing more sophisticated beliefs (Independent Knowing or Contextual Knowing). For example, an overall EBAPS score ranging between 75-100% suggests the test-taker may have beliefs aligned with Contextual Knowing, such as evidence being a requirement for understanding and that an individual is responsible for developing their own criteria for finding solutions to problems. A score closer to 75%, however, suggests that the individual may hold a mixture of sophisticated and less sophisticated beliefs.

Total EBAPS scores at the beginning of the semester (N = 257, M = 80.0 or 67%, SD = 10.7) were not statistically significantly different from scores at the end of the semester (N = 119, M = 77.4 or 65%, SD = 11.5), as summarized in Tables 2.5 and 2.6. A linear regression analysis was run to investigate the effects of EBAPS scores at the beginning of semester, sex, lecturer, course performance, and interactions between these effects on EBAPS scores at the end of the semester. The final regression equation ($R^2 = 0.427$, $F(8,60) = 5.581$, $p < 0.001$) included only the EBAPS score at the beginning of the semester as a statistically significant predictor of end-of-semester EBAPS score ($p < 0.001$, $sr^2 = 0.338$). No student characteristics led to differences in EBAPS scores at the end of the semester based on this analysis.
Difference in EBAPS scores between sexes

Comparisons between sexes were made using independent samples t-tests. Both time points were analyzed. The assessment at the end of the semester showed no differences between scores of men and women on any subscale or the total score, but the assessment given at the beginning of the semester shows a statistically significant difference in total EBAPS score ($p = 0.016$) and in the scores for Subscale 2 ($p = 0.001$). All other scores for the first time period were not different statistically. On average, women scored 67.4% on the total EBAPS and men scored 63.5%, a difference of 3.9%. The effect size of this difference, however, is small ($\eta^2 = 0.024$). Subscale 2 examines the understanding of the nature of knowing and learning. On average, women scored 65.0% on that scale while men scored 57.3%, a difference of 7.7%. The effect size of this difference is also small ($\eta^2 = 0.009$) Equal variances were assumed for all scores. Because the sample size for men was so small for the second EBAPS test (N=12) relative to the number of women (N=92), the significance of differences in results will not resolve easily statistically. A larger sample size will give a better view of differences. Total scores on the EBAPS and subscale scores did not change over time for women. Results of the independent and paired samples t-tests are summarized in Tables 2.7 and 2.8.
Table 2.5 Independent samples t-test results \((\alpha = 0.05)\) comparing Epistemological Beliefs Assessment for Physical Science (EBAPS) scores at the beginning of the semester with scores at the end of the semester.

<table>
<thead>
<tr>
<th></th>
<th>Total Score (beginning of semester, end of semester)</th>
<th>Subscale 1 (Structure of scientific knowledge)</th>
<th>Subscale 2 (Nature of knowing and learning)</th>
<th>Subscale 3 (Real-life applicability)</th>
<th>Subscale 4 (Evolving knowledge)</th>
<th>Subscale 5 (Source of ability to learn)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>257, 119</td>
<td>257, 119</td>
<td>257, 119</td>
<td>257, 119</td>
<td>257, 119</td>
<td>257, 119</td>
</tr>
<tr>
<td>p</td>
<td>0.078</td>
<td>0.074</td>
<td>0.339</td>
<td>.609</td>
<td>0.095</td>
<td>0.506</td>
</tr>
<tr>
<td>t</td>
<td>1.767</td>
<td>1.789</td>
<td>0.958</td>
<td>0.511</td>
<td>1.674</td>
<td>0.666</td>
</tr>
<tr>
<td>(\eta^2)</td>
<td>0.008</td>
<td>0.008</td>
<td>0.002</td>
<td>0.001</td>
<td>0.007</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table 2.6 Paired samples t-test results \((\alpha = 0.05)\) comparing Epistemological Beliefs Assessment for Physical Science (EBAPS) scores at the beginning of the semester with scores at the end of the semester for students who completed both assessments.

<table>
<thead>
<tr>
<th></th>
<th>Total Score (Structure of scientific knowledge)</th>
<th>Subscale 2 (Nature of knowing and learning)</th>
<th>Subscale 3 (Real-life applicability)</th>
<th>Subscale 4 (Evolving knowledge)</th>
<th>Subscale 5 (Source of ability to learn)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>p</td>
<td>0.743</td>
<td>0.856</td>
<td>0.953</td>
<td>0.561</td>
<td>0.571</td>
</tr>
<tr>
<td>t</td>
<td>0.329</td>
<td>-0.182</td>
<td>0.059</td>
<td>0.585</td>
<td>-0.568</td>
</tr>
<tr>
<td>(\eta^2)</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.004</td>
<td>0.004</td>
</tr>
</tbody>
</table>


Table 2.7 Independent samples t-test results (α = 0.05) comparing Epistemological Beliefs Assessment for Physical Science (EBAPS) scores (as percentages) of females (dummy variable = 1) and males (dummy variable = 0) at the beginning of the semester. Statistically significantly different scores are bolded.

<table>
<thead>
<tr>
<th></th>
<th>Total Score (females, males)</th>
<th>Subscale 1, Structure of scientific knowledge (females, males)</th>
<th>Subscale 2, Nature of knowing and learning (females, males)</th>
<th>Subscale 3, Real-life applicability (females, males)</th>
<th>Subscale 4, Evolving knowledge (females, males)</th>
<th>Subscale 5, Source of ability to learn (females, males)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>213, 35</td>
<td>213, 35</td>
<td>213, 35</td>
<td>213, 35</td>
<td>213, 35</td>
<td>213, 35</td>
</tr>
<tr>
<td>M</td>
<td>67.4%, 63.5%</td>
<td>59.2%, 56.8%</td>
<td>65.0%, 57.3%</td>
<td>73.5%, 74.9%</td>
<td>70.2%, 65.7%</td>
<td>78.5%, 74.3%</td>
</tr>
<tr>
<td>SD</td>
<td>8.5%, 10.1%</td>
<td>10.8%, 11.0%</td>
<td>11.8%, 15.1%</td>
<td>16.4%, 14.5%</td>
<td>17.7%, 17.6%</td>
<td>15.2%, 17.2%</td>
</tr>
<tr>
<td>p</td>
<td>0.016</td>
<td>0.202</td>
<td>0.001</td>
<td>0.622</td>
<td>0.167</td>
<td>0.137</td>
</tr>
<tr>
<td>t</td>
<td>2.422</td>
<td>1.279</td>
<td>3.421</td>
<td>-0.494</td>
<td>1.386</td>
<td>1.493</td>
</tr>
<tr>
<td>η²</td>
<td>0.024</td>
<td>0.007</td>
<td>0.009</td>
<td>0.045</td>
<td>0.008</td>
<td>0.009</td>
</tr>
</tbody>
</table>
Table 2.8 Independent samples t-test results ($\alpha = 0.05$) comparing Epistemological Beliefs Assessment for Physical Science (EBAPS) scores as percentages of females (dummy variable = 1) and males (dummy variable = 0) at the end of the semester.

<table>
<thead>
<tr>
<th></th>
<th>Total Score (females, males)</th>
<th>Subscale 1, Structure of scientific knowledge (females, males)</th>
<th>Subscale 2, Nature of knowing and learning (females, males)</th>
<th>Subscale 3, Real-life applicability (females, males)</th>
<th>Subscale 4, Evolving knowledge (females, males)</th>
<th>Subscale 5, Source of ability to learn (females, males)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>92, 12</td>
<td>92, 12</td>
<td>92, 12</td>
<td>92, 12</td>
<td>92, 12</td>
<td>213, 35</td>
</tr>
<tr>
<td>M</td>
<td>64.6%, 64.2%</td>
<td>55.9%, 60.3%</td>
<td>62.4%, 58.9%</td>
<td>72.0%, 71.6%</td>
<td>64.6%, 72.2%</td>
<td>77.5%, 71.2%</td>
</tr>
<tr>
<td>SD</td>
<td>9.3%, 8.8%</td>
<td>10.8%, 11.0%</td>
<td>14.1%, 13.5%</td>
<td>16.2%, 16.3%</td>
<td>20.2%, 19.6%</td>
<td>14.6%, 19.0%</td>
</tr>
<tr>
<td>p</td>
<td>0.890</td>
<td>0.164</td>
<td>0.416</td>
<td>0.942</td>
<td>0.219</td>
<td>0.183</td>
</tr>
<tr>
<td>t</td>
<td>0.139</td>
<td>-1.403</td>
<td>0.817</td>
<td>0.073</td>
<td>-1.236</td>
<td>1.342</td>
</tr>
<tr>
<td>$\eta^2$</td>
<td>0.024</td>
<td>0.007</td>
<td>0.009</td>
<td>0.045</td>
<td>0.008</td>
<td>0.009</td>
</tr>
</tbody>
</table>
Differences between students of Lecturer A and Lecturer B

Comparisons between participants with different lectures were made using independent samples t-test for EBAPS scores for both time periods. At the beginning of the semester, total EBAPS scores and scores for Subscales 1-4 were found to be insignificantly different. However, the scores for Subscale 5 (the source of ability to learn) were found to be significantly different at a confidence level of 95% (p = 0.007). Students of Lecturer A (N = 178) scored, on average, 76.3% on this subscale compared to an average score of 82.1% by students of lecturer B (N = 74). This difference of 5.8% has a small effect size ($\eta^2 = 0.029$), however, due to the large variance in the sample. Lecturer B reported using more active learning strategies in the classroom compared to Lecturer A. As Subscale 5 investigates the understanding of the source of one’s ability to learn, students experiencing active learning activities may be engaged in learning strategies they have not before experienced, and may begin to change how they feel about how learning involves hard work rather than innate ability. One hypothesis about the difference in beliefs about the innate ability to learn at the beginning of the semester may be explained by the activities the students were asked to do in class with Lecturer B. At the end of the semester, however, the difference in Subscale 5 scores disappeared as the score for students of Lecturer B decreased from 82% to 79%, and no other scores were found to be statistically significantly different at $\alpha = 0.05$. The results of the independent samples t-tests are presented in Tables 2.9 and 2.10.

Overall, EBAPS scores did not change from the beginning to the end of the semester, as discussed above. To determine whether this holds true when participants are separated by lecturer, paired samples t-tests were performed on scores by students who completed the EBAPS during both test times. The results of these t-tests show that Lecturer did not have an effect on
change in EBAPS score over the course of the semester at a 95% confidence level, as there were no statistically significant differences in total or subscale scores for either lecturer group, as presented in Table 2.11.
Table 2.9 Independent samples t-test results ($\alpha = 0.05$) comparing Epistemological Beliefs Assessment for Physical Science (EBAPS) scores (as percentages) of students of Lecturer A (dummy variable = 1) and Lecturer B (dummy variable = 0) at the beginning of the semester. Results for students of Lecturer A are reported first, then results for students of Lecturer B are in parentheses. Statistically significantly different scores ($\alpha=0.05$) are bolded.

<table>
<thead>
<tr>
<th></th>
<th>Total Score Lecturer A (Lecturer B)</th>
<th>Subscale 1, Structure of scientific knowledge Lecturer A (Lecturer B)</th>
<th>Subscale 2, Nature of knowing and learning Lecturer A (Lecturer B)</th>
<th>Subscale 3, Real-life applicability Lecturer A (Lecturer B)</th>
<th>Subscale 4, Evolving knowledge Lecturer A (Lecturer B)</th>
<th>Subscale 5, Source of ability to learn Lecturer A (Lecturer B)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N</strong></td>
<td>178 (74)</td>
<td>178 (74)</td>
<td>178 (74)</td>
<td>178 (74)</td>
<td>178 (74)</td>
<td>178 (74)</td>
</tr>
<tr>
<td><strong>M</strong></td>
<td>66.4% (68.0%)</td>
<td>63.8% (64.1%)</td>
<td>74.3% (72.3%)</td>
<td>68.5% (72.1%)</td>
<td></td>
<td>76.3% (82.1%)</td>
</tr>
<tr>
<td><strong>SD</strong></td>
<td>8.9% (8.8%)</td>
<td>12.1% (13.7%)</td>
<td>16.1% (16.5%)</td>
<td>17.5% (18.2%)</td>
<td></td>
<td>15.6% (14.2%)</td>
</tr>
<tr>
<td><strong>p</strong></td>
<td>0.198</td>
<td>0.816</td>
<td>0.363</td>
<td>0.145</td>
<td></td>
<td>0.007</td>
</tr>
<tr>
<td><strong>t</strong></td>
<td>-1.292</td>
<td>-0.233</td>
<td>0.912</td>
<td>-1.461</td>
<td></td>
<td>-2.721</td>
</tr>
<tr>
<td>$\eta^2$</td>
<td>0.007</td>
<td>0.000</td>
<td>0.003</td>
<td>0.008</td>
<td></td>
<td>0.029</td>
</tr>
</tbody>
</table>
Table 2.10 Independent samples t-test results ($\alpha = 0.05$) comparing Epistemological Beliefs Assessment for Physical Science (EBAPS) scores (as percentages) of students of Lecturer A (dummy variable = 1) and Lecturer B (dummy variable = 0) at the end of the semester. Results for students of Lecturer A are reported first, then results for students of Lecturer B are in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>Total Score</th>
<th>Subscale 1, Structure of scientific knowledge</th>
<th>Subscale 2, Nature of knowing and learning</th>
<th>Subscale 3, Real-life applicability</th>
<th>Subscale 4, Evolving knowledge</th>
<th>Subscale 5, Source of ability to learn</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lecturer A</td>
<td>Lecturer A</td>
<td>Lecturer A</td>
<td>Lecturer A</td>
<td>Lecturer A</td>
<td>Lecturer A</td>
</tr>
<tr>
<td></td>
<td>(Lecturer B)</td>
<td>(Lecturer B)</td>
<td>(Lecturer B)</td>
<td>(Lecturer B)</td>
<td>(Lecturer B)</td>
<td>(Lecturer B)</td>
</tr>
<tr>
<td>N</td>
<td>82 (23)</td>
<td>82 (23)</td>
<td>82 (23)</td>
<td>82 (23)</td>
<td>82 (23)</td>
<td>82 (23)</td>
</tr>
<tr>
<td>M</td>
<td>64.4% (65.2%)</td>
<td>56.0% (57.7%)</td>
<td>61.6% (63.2%)</td>
<td>72.6% (69.4%)</td>
<td>65.6% (64.1%)</td>
<td>76.3% (78.7%)</td>
</tr>
<tr>
<td>SD</td>
<td>9.2% (9.0%)</td>
<td>10.3% (10.4%)</td>
<td>14.4% (12.6%)</td>
<td>15.2% (19.2%)</td>
<td>20.1% (20.6%)</td>
<td>14.9% (16.5%)</td>
</tr>
<tr>
<td>p</td>
<td>0.724</td>
<td>0.495</td>
<td>0.625</td>
<td>0.401</td>
<td>0.718</td>
<td>0.514</td>
</tr>
<tr>
<td>t</td>
<td>-1.354</td>
<td>-0.685</td>
<td>-0.491</td>
<td>0.843</td>
<td>0.362</td>
<td>-0.655</td>
</tr>
<tr>
<td>$\eta^2$</td>
<td>0.017</td>
<td>0.005</td>
<td>0.002</td>
<td>0.007</td>
<td>0.005</td>
<td>0.004</td>
</tr>
</tbody>
</table>
Table 2.11 Paired samples t-test results ($\alpha = 0.05$) comparing Epistemological Beliefs Assessment for Physical Science (EBAPS) scores at the beginning of the semester with scores at the end of the semester for students who completed both assessments. Results for students of Lecturer A are reported first, then results for students of Lecturer B are in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>Total Score Lecturer A (Lecturer B)</th>
<th>Subscale 1 (Structure of scientific knowledge)</th>
<th>Subscale 2 (Nature of knowing and learning)</th>
<th>Subscale 3 (Real-life applicability)</th>
<th>Subscale 4 (Evolving knowledge)</th>
<th>Subscale 5 (Source of ability to learn)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>54 (16) 54 (16)</td>
<td>54 (16)</td>
<td>54 (16)</td>
<td>54 (16)</td>
<td>54 (16)</td>
<td>54 (16)</td>
</tr>
<tr>
<td>p</td>
<td>0.500 (0.691) 0.967 (0.354)</td>
<td>0.533 (0.744)</td>
<td>0.223 (0.892)</td>
<td>0.718 (0.283)</td>
<td>0.963 (0.860)</td>
<td></td>
</tr>
<tr>
<td>t</td>
<td>0.679 (0.405) -0.041 (0.956)</td>
<td>0.628 (-0.333)</td>
<td>1.234 (-0.138)</td>
<td>-0.362 (1.112)</td>
<td>0.046 (-0.180)</td>
<td></td>
</tr>
<tr>
<td>$\eta^2$</td>
<td>0.009 (0.011) 0.000 (0.057)</td>
<td>0.007 (0.007)</td>
<td>0.028 (0.001)</td>
<td>0.002 (0.076)</td>
<td>0.000 (0.002)</td>
<td></td>
</tr>
</tbody>
</table>
Correlations between EBAPS scores and course performance

Because epistemological beliefs can directly influence how students are able to comprehend topics in science, the relationship between course performance and EBAPS scores was investigated. EBAPS scores taken at the beginning of the semester were correlated with Z-scores of the first midterm of the semester, and EBAPS scores from the end of the semester were correlated with final exam Z-scores.

At the beginning of the semester, midterm 1 Z-scores correlate positively with EBAPS total score (r = 0.231, p < 0.001), Subscale 1 (r = 0.156, p = 0.014), Subscale 2 (r = 0.141, p = 0.028), and Subscale 4 (r = 0.135, p = 0.035). The correlations between midterm scores and Subscales 3 and 5 were not statistically significant at 95% confidence. At the end of the semester, EBAPS total score (r = 0.233, p = 0.017) and Subscales 1 (r = 0.298, p = 0.002) and 4 (r = 0.236, p = 0.015) correlate positively with final exam Z-scores. Results are presented in Tables 2.12 and 2.13. In both cases, real-life applicability and the source of ability to learn to do not correlate with course performance. Subscale 2 (the nature of knowing and learning), though correlating with performance at the beginning of the semester, does not correlate at the end of the semester. As discussed above, students’ scores on Subscale 2 (the nature of knowing and learning) do not change over the course of the semester. However, as the semester progresses, there ceases to be a correlation between performance and Subscale 2 score.

Table 2.12 Pearson correlations for midterm Z-scores and EBAPS scores at the beginning of the semester (N=245). Statistically significant correlations (α=0.05) are bolded.

<table>
<thead>
<tr>
<th></th>
<th>Total score</th>
<th>Subscale 1</th>
<th>Subscale 2</th>
<th>Subscale 3</th>
<th>Subscale 4</th>
<th>Subscale 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Midterm 1 Z-score</td>
<td>0.231</td>
<td>0.156</td>
<td>0.141</td>
<td>0.106</td>
<td><strong>0.135</strong></td>
<td>0.113</td>
</tr>
<tr>
<td>p (2-tailed)</td>
<td>&lt;0.001</td>
<td>0.014</td>
<td>0.028</td>
<td>0.099</td>
<td>0.035</td>
<td>0.076</td>
</tr>
</tbody>
</table>
Table 2.13 Pearson correlations for final Z-scores and EBAPS scores at the end of the semester (N=105). Statistically significant correlations (α=0.05) are bolded.

<table>
<thead>
<tr>
<th></th>
<th>Total score</th>
<th>Subscale 1</th>
<th>Subscale 2</th>
<th>Subscale 3</th>
<th>Subscale 4</th>
<th>Subscale 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final Z-score</td>
<td><strong>0.233</strong></td>
<td><strong>0.298</strong></td>
<td>0.135</td>
<td>0.084</td>
<td><strong>0.236</strong></td>
<td>-0.017</td>
</tr>
<tr>
<td>p (2-tailed)</td>
<td>0.017</td>
<td>0.002</td>
<td>0.169</td>
<td>0.392</td>
<td>0.015</td>
<td>0.865</td>
</tr>
</tbody>
</table>

Analysis using linear regression

Linear regression analysis was employed to get a clearer understanding of what factors may affect epistemological beliefs (as measures quantitatively by the EBAPS). From the exploratory analyses, lecturer, sex, and time do not appear to affect EBAPS total or subscores, but course performance was found to be correlated with EBAPS total scores and some subscores. The relationships between the main effects of lecturer, sex, and course performance, and the interactions between these main effects, were investigated.

Predictors of EBAPS scores at the beginning of the semester

A statistical regression model was run to investigate the main effects of sex, lecturer, and midterm 1 Z-score, and interactions between these effects. Both standardized (β) and unstandardized (B) regression coefficients are reported in all cases where regression analysis was employed. The focus is on the unstandardized coefficients as variables containing different scales and means (the measures of course performance) have already been standardized as Z-scores. Other variables of sex and lecturer were categorical and dummy coded. Squared semipartial correlations (sr²) for predictor variables were reported. These values represent the portion of the overall variance of the outcome measure that can be uniquely predicted by the predictor variable. For a particular predictor, sr² can be interpreted as an effect size when all
other variables are statistically controlled. The semipartial correlation ($sr$) was used to apply Cohens’s guidelines.

The model tested is represented by the equation:

$$EBAPS_i = b_0 + b_1(sex) + b_2(lecturer) + b_3(midterm \ 1 \ Z \ - \ score) + b_4(sex \ * \ lecturer)$$
$$+ b_5(sex \ * \ midterm \ 1 \ Z \ - \ score) + b_6(lecturer \ * \ midterm \ 1 \ Z \ - \ score)$$
$$+ b_6(sex \ * \ lecturer \ * \ midterm \ 1 \ Z \ - \ score)$$

The order of entry into the regression analysis was decided based on which predictor gave the largest increase in $R^2$ when added to the regression equation. At the beginning of the analysis, there were no predictors, and predictors were added in steps. If at a given step, however, a predictor that had been added previously no longer strongly contributed to $R^2$ of the model, that predictor was dropped from the model (Warner, 2013, p 560-561).

The results of the statistical regression show that only the main effects of midterm score and sex were statistically significant predictors of total EBAPS scores at the beginning of the semester, as summarized in Table 2.14. No other effects included in this study contributed significantly to the $R^2$ value of the model. Overall, approximately 6.8% of the variance of total EBAPS scores could be accounted for by the regression (adjusted $R^2 = 0.068$, $F(2,240) = 9.832$, $p < 0.001$). On its own, course performance was able to explain 5% ($sr^2 = 0.050$) of the variance in EBAPS scores. When sex was entered into the model, 2.6% ($sr^2 = 0.026$) of the variance of scores were explained by this main effect. The squared semipartial for course performance suggests a medium effect size when sex is statistically controlled. The $sr^2$ value for sex suggests a small effect size for this predictor when course performance is statistically controlled. The regression model for predicting EBAPS scores at the beginning of the semester was
\[ EBAPS_i = 75.9 + 2.5(midterm\ 1\ Z\ -\ score) + 4.8(sex) \]

On average, with all effects held constant, students scored an average of 75.9 points out of 120 on the EBAPS at the beginning of the semester. Women, on average, scored 4.8 points higher than men, controlling for course performance. Students who had scores one standard deviation above the mean on their first midterm were predicted to score 2.5 points higher on the EBAPS than those who had mean midterm scores, controlling for sex.

Table 2.14 Statistical regression analysis results for predictors of EBAPS total score at the beginning of the semester, variables added stepwise to the equation. (Total points for EBAPS is 120.)

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>Standard error</th>
<th>( \beta )</th>
<th>t</th>
<th>( sr^2 )</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant (b_0)</td>
<td>80.0</td>
<td>0.661</td>
<td>121.1</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Midterm 1 Z - score</td>
<td>2.451</td>
<td>0.686</td>
<td>0.224</td>
<td>3.571</td>
<td>0.050</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>( R^2 = 0.050 ), Adjusted ( R^2 = 0.046 ), R = 0.224, F(1, 240) = 12.751, p &lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant (b_0)</td>
<td>75.9</td>
<td>1.741</td>
<td>43.585</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Midterm 1 Z - score</td>
<td>2.496</td>
<td>0.679</td>
<td>0.228</td>
<td>3.677</td>
<td>0.053</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sex</td>
<td>4.820</td>
<td>1.874</td>
<td>0.160</td>
<td>2.572</td>
<td>0.026</td>
<td>0.011</td>
</tr>
<tr>
<td>( R^2 = 0.076 ), Adjusted ( R^2 = 0.068 ), R = 0.275, F(2, 240) = 9.832, p &lt; 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Predictors of EBAPS scores at the end of the semester

Correlations were found between course performance at the end of the semester (final exam Z-score) and EBAPS scores, but no significant correlations were found between EBAPS scores and the main effects of sex and lecturer. A stepwise statistical linear regression was performed with the main effects of sex, lecturer, course performance (final exam Z-score), and with interactions between these main effects. The regression model tested was

$$EBAPS_i = b_0 + b_1(\text{sex}) + b_2(\text{lecturer}) + b_3(\text{final } Z - \text{score}) + b_4(\text{sex} * \text{lecturer})$$

$$+ b_5(\text{sex} * \text{final } Z - \text{score}) + b_6(\text{lecture} * \text{final } Z - \text{score})$$

$$+ b_6(\text{sex} * \text{lecturer} * \text{final } Z - \text{score})$$

The same procedure for the regression for beginning of semester EBAPS score prediction was employed for this analysis.

The results of the regression analysis for EBAPS scores at the end of the semester are provided in Table 2.15 The model explained 4.5% of the variance in EBAPS scores (adjusted $R^2 = 0.045$, $F(1,102) = 5.486, p = 0.017$). At the end of the semester, the only statistically significant predictor for overall EBAPS score was found to be course performance ($p = 0.017$). At the beginning of the semester, women were predicted to score an average of five points higher on the EBAPS than men, but this difference no longer appears in the scores at the end of the semester. The squared semipartial of the course performance outcome ($sr^2 = 0.054$) suggests a medium effect size. The final model was:

$$EBAPS_i = 77.7 + 2.97(\text{final } Z - \text{score})$$
Table 2.15 Statistical regression analysis results for predictors of EBAPS total score at the end of the semester, variables added stepwise to the equation. (Total points for EBAPS is 120.)

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>Standard error</th>
<th>β</th>
<th>t</th>
<th>sr²</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant (b₀)</td>
<td>77.702</td>
<td>1.059</td>
<td>73.347</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final Z-score</td>
<td>2.971</td>
<td>1.229</td>
<td>0.233</td>
<td>2.418</td>
<td>0.054</td>
<td>0.017</td>
</tr>
</tbody>
</table>

R² = 0.054, Adjusted R² = 0.045, R = 0.233, F(1, 102) = 5.846, p = 0.017

Effect of lecturer on Subscale 5 scores

As expected from the exploratory analysis, the lecturer the student had for general chemistry class did not predict level of epistemological beliefs. However, the two lecturers did report employing different teaching methods during the semester, so it would be valuable to see whether an effect of lecturer did become significant as students may have had different experiences with learning over the course of the semester. In particular, it would be of interest to evaluate whether the lecturer had an effect on Subscale 5 (the source of one’s ability to learn) of the EBAPS as there were statistically different scores on this subscale at the beginning of the semester.

The stepwise statistical regression analysis was run with scores Subscale 5 as the dependent variables (measured at the beginning of the semester, during the third week of instruction). The results of the analysis (presented in Table 2.16) gave a model that explains 4.3% of the variation in Subscale 5 scores (adjusted R² = 0.043, F(2,240) = 6.413, p = 0.002). Two effects were found to contribute significantly to the model: lecturer (p = 0.001, sr² = 0.048) and the interaction between sex and lecturer (p = 0.018, sr² = 0.023). The model for predicting Subscale 5 scores is
Subscale 5 \_t = 16.38 - 2.6(Lecturer) + 1.7(sex \ast Lecturer)

The negative correlation between lecturer and subscale score suggests that students of Lecturer A (dummy value =1) will score, on average, 2.6 points lower on Subscale 5 than students of Lecturer B, which is a difference of 13% on the 20-point scale. The squared semipartial of the effect of Lecturer ($sr^2 = 0.048$) suggests a small-to-medium effect size when the effect of the interaction between sex and lecturer was statistically controlled. The interaction term is more difficult to interpret, as both sex and lecturer were dummy variables. If both dummy variables were 1 (female students of Lecturer A), scores on Subscale 5 would be, on average, 1.7 points higher than an individual with any of the dummy variables equaling 0 (male students or any student of Lecturer B). One possible interpretation of the interaction term is that Lecturer A’s teaching style differentially affected students based on sex. The effect size, based on the squared semipartial of the interaction ($sr^2 = 0.022$), is suggested to be small. When the same effects were tested to predict Subscale 5 outcomes at the end of the semester, none were statistically significant predictors (Lecturer p = 0.654, sex p = 0.925).

Effect of course performance on subscale scores at the beginning of the semester

Stepwise statistical regression analyses were performed with each subscale score as the outcome variable. Three of the five subscales were found to have course performance (as measured by score on a midterm) to be a statistically significant predictors, either in the form of a main effect or as part of an interaction effect. Results of these regression analysis are presented in Table 2.17.

Subscale 1 measured beliefs about the structure of scientific knowledge. The resulting model from the regression analysis predicted 3.0% of variance of Subscale 1 scores (adjusted $R^2$
The only variable that emerged as a statistically significant predictor of the outcome score was the interaction between course performance and sex \((p = 0.004, \text{sr}^2 = 0.033)\). On average, students scored 23.5 points out of 40 on this subscale. The positive coefficient of 0.901 for sex (dummy score for females = 1) means that female students who score one standard deviation above the mean on the midterm would be predicted to score one point higher on this subscale than the average student. The final model for predicting Subscale 1 scores was:

Subscale 1\(_i\) = 23.5 + 0.9(Midterm 1 Z-score * sex).

Subscale 2 measures beliefs about the nature of knowing and learning. The statistically significant predictors found for outcomes of this scale were sex and course performance, and the model predicted 5.9% of variance of outcome scores (adjusted \(R^2 = 0.059, F(2,240) = 8.550, p < 0.001\)). On average, students scored 18.3 points out of 32 on this subscale. The coefficient for sex was 2.493, so on average, women scored 2.5 points higher on this scale (out of a total of 32 points) than men, controlling for course performance \((p = 0.001, \text{sr}^2 = 0.046)\). The effect size of sex was small-medium when course performance was statistically controlled, based on the squared semipartial. Course performance was also found to be a statistically significant predictor \((p = 0.018, \text{sr}^2 = 0.023)\), but had a small effect size when sex was statistically controlled, based on the squared semipartial. Students scoring one standard deviation above the mean on the midterm are predicted to score approximately half a point higher than the average student. The gain here is modest, so high performers on the midterm are not predicted to have EBAPS scores much higher than the average. The effect size for course performance is small. The final model for predicting Subscale 2 scores was:
Finally, Subscale 4, which measures beliefs about the tendency for knowledge to evolve, was analyzed. Course performance was the only predictor which significantly contributed to the model, which could predict 1.5% of variance of Subscale 4 scores (out of 12 points) (adjusted $R^2 = 0.015$, $F(1,241) = 4.637$, $p = 0.032$). On average, students score 8.3 points on this subscale when course performance was statistically controlled. Students who scored one standard deviation above the mean on the midterm were predicted to score approximately 0.3 points higher on Subscale 4 than students who scored the mean on the midterm ($p = 0.032$, $sr^2 = 0.019$), but this predictor had only a small effect size. Again, any gains on the Subscale predicted by course performance were small. The final model for predicting Subscale 4 scores was:

$$Subscale\ 4_i = 8.3 + 0.3(Midterm\ 1\ Z - score \times sex).$$
Table 2.16 Statistical regression analysis results for predictors of Subscale 5 at the beginning of the semester, variables added stepwise to the equation. (Total points of Subscale 5 is 20.)

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>Standard error</th>
<th>β</th>
<th>t</th>
<th>sr²</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant (b₀)</td>
<td>16.380</td>
<td>0.359</td>
<td>45.631</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Lecturer (Lecturer A = 1, Lecturer B = 0)</td>
<td>-1.130</td>
<td>0.427</td>
<td>-0.168</td>
<td>-2.649</td>
<td>0.028</td>
<td>0.009</td>
</tr>
<tr>
<td>R² = 0.028, Adjusted R² = 0.024, R = 0.168, F(1, 241) = 7.017, p = 0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>Standard error</th>
<th>β</th>
<th>t</th>
<th>sr²</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant (b₀)</td>
<td>16.380</td>
<td>0.356</td>
<td>46.071</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Lecturer (Lecturer A = 1, Lecturer B = 0)</td>
<td>-2.630</td>
<td>0.758</td>
<td>-0.391</td>
<td>-3.468</td>
<td>0.048</td>
<td>0.001</td>
</tr>
<tr>
<td>Sex * Lecturer (female = 1, male = 0)</td>
<td>1.697</td>
<td>0.713</td>
<td>0.269</td>
<td>2.382</td>
<td>0.022</td>
<td>0.018</td>
</tr>
<tr>
<td>R² = 0.051, Adjusted R² = 0.043, R = 0.225, F(2, 240) = 6.413, p = 0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.17 Statistical regression analysis results for predictors of each subscale at the beginning of the semester, variables added stepwise to the equation.

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>Standard error</th>
<th>β</th>
<th>t</th>
<th>sr²</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subscale 1 (total of 40 points)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant (b₀)</td>
<td>23.534</td>
<td>0.272</td>
<td>86.61</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midterm 1 Z-score x Sex (1=female, 0=male)</td>
<td>0.901</td>
<td>0.312</td>
<td>0.183</td>
<td>2.891</td>
<td>0.033</td>
<td>0.004</td>
</tr>
<tr>
<td><strong>R² = 0.034, Adjusted R² = 0.030, R = 0.183, F(1, 241) = 8.359, p = 0.004</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Subscale 2 (total of 32 points)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant (b₀)</td>
<td>18.256</td>
<td>0.673</td>
<td>27.120</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sex (female = 1, male = 0)</td>
<td>2.493</td>
<td>0.725</td>
<td>0.215</td>
<td>3.440</td>
<td>0.046</td>
<td>0.001</td>
</tr>
<tr>
<td>Midterm 1 Z-score</td>
<td>0.625</td>
<td>0.262</td>
<td>0.149</td>
<td>2.383</td>
<td>0.023</td>
<td>0.018</td>
</tr>
<tr>
<td><strong>R² = 0.067, Adjusted R² = 0.059, R = 0.258, F(2, 240) = 8.550, p &lt; 0.001</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Subscale 4 (total of 12 points)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant (b₀)</td>
<td>8.334</td>
<td>0.135</td>
<td>61.602</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midterm 1 Z-score</td>
<td>0.303</td>
<td>0.141</td>
<td>2.153</td>
<td>0.019</td>
<td>0.032</td>
<td></td>
</tr>
<tr>
<td><strong>R² = 0.019, Adjusted R² = 0.015, R = 0.137, F(1, 241) = 4.637, p = 0.032</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Discussion and conclusions

1. What level of sophistication of epistemological beliefs do general chemistry students hold?

Scores on the EBAPS give an indication of the level of sophistication of students’ epistemological beliefs. The population in this study scored values comparable to those of college physics students (Marx et al., 2004) and high school physics students (Elby, 2001). The scores, when applied to Baxter Magolda’s Model of Epistemological Reflection, suggest that general chemistry students have moderately sophisticated views about the nature of knowledge and learning. A total EBAPS score of 67% suggests that, on average, students hold beliefs straddling Transitional Knowing and Independent Knowing (Baxter Magolda, 2004). In that study, she found that about 53% of college students used the pattern of Transitional Knowing in their sophomore year, and 16% of senior college students were using the pattern of independent knowing. Many students at the Transitional Knowing level understand science knowledge as concrete facts that have been discovered and analyzed by experts, and only in some cases can these facts be disproven or only applicable in certain cases. Students are beginning to understand that learning is a student-centered activity that relies on hard work and observations of the world (rather than a passive activity where learning relies on experts imparting “true” knowledge). When knowledge is assessed, students with moderate sophistication levels of EB expect the instructors to be able to evaluate how well the material is understood, but these students do not take into account that they themselves should have the ability to assess their own knowledge. These individuals still mainly rely on experts to validate knowledge, though they are beginning to understand that they can learn from (and contribute to) knowledge of their peers.

Models of epistemological beliefs include different dimensions of beliefs (Schommer et al., 1992; Hofer & Pintrich, 1997; Schommer-Aikins, 2004), and development of the EBAPS
questionnaire included subscales to measure levels of beliefs of five dimensions. These dimensions include the structure of scientific knowledge, the nature of knowing and learning, real-life applicability of science, the source of an individual’s ability to learn, and the evolution of knowledge. Students’ responses to items that belong to a subscale are meant to measure the student’s beliefs about that dimension, but it can be difficult to tease out a student’s beliefs apart from their goals or expectations (diSessa, 1985; Hammer & Elby, 2001). Interpreting scores on subscales as sophistication of just beliefs can lead to some misunderstandings of a student’s actual belief system, and so this limitation of the instrument used must be kept in mind when analyzing results.

Three of the five subscales had average scores between 58-69%: Structure of scientific knowledge (58.8%), Nature of knowing and learning (63.8%), and Evolving knowledge (69.2%). The scores on these subscales suggest a moderate level of sophistication in terms of epistemological beliefs, which would align most closely with Baxter Magolda’s Transitional Knowing patterns (2004). Students are uncomfortable with uncertainty of knowledge, and expect there to be “right” and “wrong” answers about most concepts, though in some cases it is accepted that rather than a concrete, correct answer, the appropriate response to a question might require more speculation. Students in this population understand that learning requires more than rote memorization in some cases, but that for science many facts are just facts that must be memorized. This is in contrast to more sophisticated views held by experts in chemistry who understand that “knowing” a concept requires integration of personal experience, evidence, and understanding of other related concepts (Neber & Schommer-Aikins, 2002; Mohamed & El-Habbal, 2013).
The two subscales with the highest scores for this population were Real-life Applicability (73.8%) and Source of Ability to Learn (78.0%). On average, the students in this population believe that science is happening and relevant outside of the academic environment. This is encouraging to see as much work is being done to engage students in learning about science using real-world examples (Assaraf & Orion, 2005; Smith & Morgan, 2016; Yoon et al., 2018). These students also hold some understanding that scientific thinking is not merely an exercise for science classes, but that thinking critically about how ideas are presented by the media is important. Though still holding the knowledge of experts of higher validity than their own knowledge, these students appear to be on the way to using their own knowledge to make decisions and judgments about their observations.

2. To what degree do levels of EB correlate with course performance, sex, and instructor?

The influence of course performance was investigated through linear regression analysis, and it was found that there is a positive correlation between course performance as measured by a midterm or final exam and overall EBAPS score. In addition, sex was a significant predictor for EBAPS scores at the beginning of the semester. The two models for predicting total EBAPS scores are:

Beginning of semester:

\[ EBAPS_i = 75.9 + 2.5(midterm\ Z\ -\ score) + 4.8(sex) \]

End of semester:

\[ EBAPS_i = 77.7 + 2.97(final\ Z\ -\ score) \]
At both time points, course performance correlated positively with EBAPS scores, predicting that students who had greater grasp of the course materials held slightly more sophisticated epistemological beliefs overall. The effect size of course performance as a predictor was medium in both cases ($sr^2_{\text{beginning of semester}} = 0.053$, $sr^2_{\text{end of semester}} = 0.054$) when effects of other predictors (sex at the beginning of the semester) were statistically controlled. Previous studies have found higher academic achievement for those students more developed (or more sophisticated) epistemological beliefs (Schommer, 1993b; Deryakulu, 2002; Mert & Bulut, 2006). However, the differences in EBAPS scores between high and low performers were modest in this study – students scoring one standard deviation above the mean on the performance assessment were predicted to only score approximately 3 points higher on the EBAPS out of a total of 120 points, a 2.5% difference. An explanation for this modest increase in score is that students do not often come across more nuanced science topics. In these courses, students have very little experience with developing their own understanding of concepts through life experiences, and so beliefs about the importance of student-led learning and acquisition of knowledge may have little influence on course performance at this level.

At the beginning of the semester, females were found to score an average of five points higher on the EBAPS tests than men ($p = 0.011$, $sr^2 = 0.026$). Differences in epistemological beliefs between genders has been investigated since the 1980’s. Belenky et al.’s study of women (1986) found that women tend to focus on receiving information rather than mastering it, and that some women would distrust logic and analysis. These women would hold more naïve views about the importance of a personal analysis of observations to construct knowledge. Baxter Magolda (1992, p 73) found that women more often will focus on listening to knowledge to learn, and will reach out to peers when confronted with a confusing or new concept, whereas
men more often will become actively involved in classroom activities to learn, but will focus on their own understanding and perspectives when a new idea is introduced. These activities can occur amongst groups of the same sophistication level of belief, but the ways in which knowledge is gathered can differ. In the present study, however, it was found that women hold slightly more sophisticated epistemological beliefs overall than men, though the difference is small and is only statistically significant in the regression model at the beginning of the semester.

Scores on individual subscales were analyzed as outcomes to determine what variables measured in this study would predict sophistication levels of five different dimensions of epistemological beliefs. The first subscale measured sophistication about beliefs regarding the structure of scientific knowledge, such as how fragmented ideas in science are, or whether many scientific concepts work together as a whole to describe phenomena. The results of the regression analysis suggest higher achieving students are predicted to score higher on this subscale than average students, but that effect only occurs when the students are female. However, any effect the interaction between sex and course performance has on beliefs about this subscale is small. It is estimated that female students scoring one standard deviation above the mean in a course performance measure will only gain an additional one point (out of 40) \((sr^2 = 0.033)\) on this subscale, and so this interaction does not have any real significant effect on the magnitude of predicted epistemological beliefs about the structure of knowledge.

Analysis of the subscale that measures understanding of the nature of learning found that women, on average, are predicted to score 2.5 points higher on the 32-point scale than men \((p = 0.001, sr^2 = 0.046)\). The mean average score predicted for this subscale controlling for sex was 18.3 points (or 57%), suggesting that this population had beliefs more consistent with Absolute Knowing of Transitional Knowing, where learning happens mainly by passively absorbing
information and knowledge from an expert (Baxter Magolda 2004). However, at the beginning of the semester, the average difference of 7.5% scored by women suggests that there is movement toward an understanding that the individual has responsibility for his or her own attainment of knowledge. The implication of this is that during the course of the semester females have experiences that contribute to a small regression in their beliefs about learning as an active process. As mentioned previously, this may have to do with the structure of the general chemistry classroom in this population, where many topics are introduced by the lecturer in the form of facts. However, no difference between men and women was found for scores on this subscale.

Beliefs about evolution of scientific knowledge were measured by Subscale 4, and the regression results produced a model with course performance as a statistically significant predictor. The magnitude of points gained by high performers, however, was very small (B = 0.303, p = 0.032, sr² = 0.019), and so course performance has no meaningful effect on the predicted score of this subscale.

The final dimension of epistemological beliefs is of the source of an individual’s ability to learn, or whether people have an innate and fixed natural ability or whether most people can learn through hard work. Because students in this study were enrolled in courses with one of two lecturers, any difference between lecturer would give insight into how classroom environment may affect self-efficacy and beliefs. Lecturer A reported a more traditional style of teaching, whereas Lecturer B employed more process-oriented learning activities. The statistically significant regression model for predicting scores on this subscale included lecturer and the interaction between sex and lecturer as predictors (adjusted R² = 0.043, F(2,240) = 6.413, p = 0.002). Students in the more traditional lecture sections were predicted, on average, to score 2.6
points lower on the subscale (out of a total of 20 points, or 13%) than the average student at the beginning of the semester.

Studies have shown that approaches to teaching and learning have an effect on epistemological beliefs (Cano, 2005; Marx et al., 2004). The study by Marx et al. found that students in active learning-centered physics classrooms scored higher on the overall EBAPS, as well as on the final subscale, than students in a traditional classroom environment. An environment that encourages effort by the individual to gain knowledge rather than receiving knowledge passively through a lecture appears to influence how students think about an individual’s ability to learn. By experiencing learning through working through observations and experiences in the classroom, a student who may not believe he or she has an innate knack for science may begin to understand that innate ability is not required for learning science. However, at the end of the semester in this study, students enrolled in Lecturer B’s courses scored lower on this subscale than they did at the beginning of the semester, and there no longer existed a statistically significant differences in beliefs about source of ability to learn between lecturers. A possible explanation is that students in Lecturer B’s course may have experienced difficulty with the lecture style over the course of the semester. Difficult topics that didn’t come as easily to students may have an effect on whether those students believe that they required an inherent ability in order to fully learn the concept.

Another statistically significant predictor of the final subscale score was the interaction between sex and lecturer. The positive correlation with this interaction suggests that when both dummy variables equal 1 (sex = female, lecturer = Lecturer A), the subscale score would be predicted to be 1.7 points higher than the average score of 16.4 (out of 20). The interaction here implies that any differences between the environments of the classrooms of Lecturer A and
Lecturer B can be partially made up for when the student is a female. That is, overall, students of Lecturer B, on average, score higher on the final subscale. However, female students of Lecturer A score, on average, higher than male students of Lecturer A on the subscale by 1.7 points. This interaction may imply that the environment in Lecturer A’s classroom may foster independent responsibility for learning more amongst women than men.

3. How do epistemological beliefs change over the course of a semester?

Based on the lack of differences found in EBAPS scores at the beginning and end of the semester in this study, it can be concluded that overall beliefs about knowledge and the nature of learning do not change for general chemistry students (66% of whom are first-year college students). Total scores on the EBAPS and scores on each of the five subscales are unchanged over the semester. Students experience a lot of changes in their first semester at a university, some involving study and learning strategies, which are correlated with epistemological beliefs (Perry, 1970; Kitchner et al., 1983; Kitchner et al., 1989; Marx et al., 2004; Richter & Schmid, 2009)

Previous studies have shown that EB do mature over time (Perry, 1970; Schommer, 1993b; Cano, 2005), but the results of this study suggest that those changes require more than one semester in an introductory college science course. The short timescale of this study limits any inferences about the effect of time on epistemological belief maturity.

There was found to be a statistically significant difference in Subscale 5 scores at the beginning of the semester between lecturers, but by the end of the semester no significant change was detected. It may be that in the early class periods in the semester, students in class with Lecturer B were introduced to a new active learning environment, and this may have affected
beliefs about how anyone can learn with hard work, but by the end of the semester the students may have reverted to a less sophisticated understanding of learning, perhaps due to fatigue from the class or discouragement. Another explanation is that the assessments or exams in the classroom were not reflective of students’ epistemological belief systems about one’s source of ability to learn.

Classroom implications

The epistemological beliefs of experts are much more sophisticated than those of novices in the sciences (Mohamed & El-Habbal, 2013). The question about how these beliefs develop in the sciences is not well understood or studied, but it is a topic that instructors must consider if they want their students to be successful in ways of scientific thinking. Graduate students are expected to understand the importance of interpretation of observations to develop, expand on, or even to change knowledge about scientific concepts. However, development of these ideas are not explicit goals of undergraduate chemistry courses.

This study was the first to employ a quantitative survey to assess epistemological beliefs of chemistry students. The results imply that, although first semester general chemistry courses often provide new experiences to students, the epistemological beliefs they have entering the class do not change over the course of that one semester. There is some evidence, though, that beliefs about an individual’s source of ability to learn may be affected by the classroom environment to a small degree. If an instructor’s goal is to help prepare students for future chemistry or scientific courses or careers, adopting an active learning environment may be one way to do this.

Correlations between course performance and scores on the EBAPS survey were found. An implication is that by fostering more sophisticated beliefs about active rather than passive
learning, and fostering beliefs about using one’s own experiences to learn about a topic can help students make gains on chemistry assessments.

**Limitations of study**

Because this study was only performed over the course of a single semester, any changes in epistemological beliefs that chemistry students may experience were most probably not observed. Changes likely occur over a longer period of time, and so a longer longitudinal study, or a study that investigates students at several points in the undergraduate curriculum, would be more likely to uncover changes in beliefs.

A future study that focuses on teaching methods in general chemistry could be done to investigate the finding that scores on the subscale about learning ability differ by lecturer. An active learning environment may foster an understanding that individuals have the ability to learn based on hard work rather than innate ability, but more research should be done to determine whether lecture type really does influence epistemological beliefs, and to what degree.

Also, previous studies have shown that some experiences may affect epistemological belief growth. For example, students who interacted with faculty mentors during their undergraduate career were found to have more sophisticated understanding of the scientific process and the importance of using empirical evidence to understand and gain knowledge about scientific phenomena (Hunter *et al.*, 2006). This study did not include questions about student activities or experiences. It would be very insightful to see what sorts of study strategies, research or outreach activities, and interactions with peers or faculty members can influence growth of epistemological beliefs.
CHAPTER 3

DEVELOPMENT OF A TEXTBOOK USE SURVEY

Student characteristics such as reading comprehension skill, prior knowledge, and epistemological beliefs can all influence the decisions the student makes when studying. As noted in chapter 1, student prior knowledge can affect how the student can comprehend a text, and certain text characteristics can differentially aid students with varying levels of prior knowledge. Currently, there are no surveys that can explore how or why students will turn to textbooks as a study resource in chemistry.

Previous studies have investigated the features of textbooks that students find useful (Smith & Jacobs, 2003; Dávila & Talanquer, 2010; Smith, B. L. et al., 2010) or how textbooks can be improved (Gillespie, 1997; Russo, 1998; Taşdelen & Köseoğlu 2008), but few have investigated how students interact with the textbook or why they choose to use a textbook in chemistry (Smith & Jacobs, 2003; Robinson et al., 2014). This may be due, in part, to a lack of textbook use assessment tools.

The goal of this work was to develop a questionnaire that could probe how students utilize textbook materials in general chemistry, why students choose to (or not) use a textbook to study, and student attitudes toward textbooks. The specific aims of this work were:

1. To design and evaluate a survey to elucidate student textbook usage for introductory chemistry courses.

2. To use student results of the survey to validate the survey.
Methodology

General experiment design

An exploratory sequential mixed methods design was employed in this study. Mixed methods designs are often used so one method can compensate for the weaknesses of the other. A criticism of quantitative designs is that the voice of the participants is not heard, and the settings and context of the participants are not strongly conveyed in the data (Creswell & Plano Clark, 2007). Qualitative research allows for participants to voice their ideas or opinions in their own words rather than being constrained by options set forth by the researcher. However, it is up to the researcher to interpret the participants’ words, and there is room for personal bias to enter in during that interpretation. Another criticism of qualitative work is that responses from individuals can be very context-specific, and so it can be difficult to generalize findings (Creswell & Plano Clark, 2007). Both qualitative and quantitative techniques are employed in this study to produce a quantitative instrument that can assess student textbook usage. An outline of the experiment design is shown in Figure 3.1.

At the beginning of the semester, volunteers were asked to participate in a semistructured interview. The responses from these interviewed aided development of the first version of the textbook survey. During the interview a textbook was available for the participants’ and the interviewer’s referral. The goal of the interview was to determine what types of questions to ask on the survey, the language students use to describe textbook features, and the language to use in questions so the intent of the questions were clear. The notes taken by the researcher during the interviews informed how to word items and answer options on the first version of the textbook survey.
On the survey, some question items allowed for only free-response answers, some only included multiple choice options, and some included both free-response and multiple choice options for answers. Results of the first survey were analyzed both quantitatively and qualitatively. Free response answers to questions were coded so that multiple choice options could be developed. Questions that only included multiple choice answer options were quantitatively analyzed to determine whether any answer options were redundant. Questions that included both free-response and multiple choice options were analyzed to find redundancies, to find new answer options, or to re-word answer options so the intent of the answer was more clear. The next version of the survey was then developed based on the analyses of the first survey version and disseminated. The procedure was replicated two more times to produce the final textbook use survey.

The quantitative results helped determine whether redundancies occurred, or whether answer options may have been unnecessary or misunderstood (if certain answer options were chosen with very low frequency). The qualitative results helped clarify the question items and the answer options.
Figure 3.1 Sequential mixed methods study design for the development of the textbook use survey.

Participants

Student demographics are presented in Table 3.1. Approximately 65% of students who participated in taking any of the three versions of the textbook use survey were female. First-year students comprised 55% of the sample, 33% were sophomores, 8% were juniors, and 4% were seniors. The most represented major in this sample (18% of the populations) was Biomedical Science: Medical and Veterinary Sciences. 244 students completed the first version of the textbook use survey, 208 students completed the second version, and 269 students completed the third version of the survey.
Table 3.1 Demographic statistics for participants.

<table>
<thead>
<tr>
<th></th>
<th>Sex</th>
<th>Class standing</th>
<th>Academic major</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>98%</td>
<td>86%</td>
<td>84%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>65% Female</td>
<td>17.6% Biomedical Science: Medical and Veterinary Sciences</td>
</tr>
<tr>
<td></td>
<td></td>
<td>35% Male</td>
<td>16.5% Biology</td>
</tr>
<tr>
<td></td>
<td></td>
<td>55.1% First-years</td>
<td>12.0% Zoology</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>32.7% Sophomores</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8.3% Juniors</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.9% Seniors</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a: percentage of students for whom data were available

Qualitative methods

Because the textbook use survey was intended to be taken by students, it was important that the survey was designed around a student’s experience with textbooks, and that the language used in the survey would be recognizable to students. Before development of the first version of the survey, two students were interviewed using semistructured interview protocol, with questions that were designed to elucidate reasons behind using a textbook and certain textbook features as well as language to be included in the survey that would ring true to the survey takers. After each of the three surveys were disseminated, the free response answers provided by
students were coded in order to determine multiple choice options to be included in the next version of the survey.

Interviews

Two students were interviewed at the beginning of the semester during the time they were learning about chemical nomenclature. They were provided a textbook to use as a reference when answering questions. The prompts given to students were novel to this study. The interview prompts were:

1. In the last week, what resources did you use while studying for chemistry? Why did you choose to use those resources?
2. If I were to ask you a question about how to name the molecule SH₂, how would you find help from a textbook? (Students were given a textbook with the first page of the nomenclature chapter marked.)
3. Can you flip through this chapter and tell me what parts in the book you are drawn to while studying? (Students were given a textbook with the first page of the stoichiometry chapter marked.)
4. Do you have access to a textbook? How do you have access to it? When would you turn to the textbook to study?
5. If you were studying for an exam or quiz, what parts of the textbook would you find most useful?
6. When you study using a textbook, what do you do?

The students were allowed at any point to interject or change topic. They were also allowed to flip through the textbook at any time. Depending on the students’ responses they may have been asked for further explanation or clarification. Each interview lasted approximately 30 minutes and notes were taken by the interviewer during and after each interview.
One participant (Student A) was a third-year female college student majoring in biomedical sciences, enrolled in the first semester course of a two-semester traditional general chemistry sequence. This was her first time taking a general chemistry course in college. The other participant (Student B) was a first-year female college student majoring in biology, also enrolled in the first semester course of a two-semester traditional general chemistry sequence. These participants responded to an email sent out to all general chemistry lecture sections asking for participants for a short interview about study habits. The participants were given a consent form to sign that included comments about confidentiality. A copy of the consent form can be found in Appendix D.

Survey responses

Surveys were disseminated online through the university sponsored Qualtrics survey software. The consent form was included as the first item on the survey. Students were given unlimited time to respond to the 10-15 items on the survey. All items were forced response, with some requiring free response answers while others had multiple choice options. The first survey was disseminated to students in the second semester course of a two-semester traditional general chemistry sequence one week after the first midterm exam. The second version of the survey was disseminated to students in the same course midway between their second and third (of three) midterm exams. The third version of the survey was disseminated one week prior to the final exam in the same course.

The codes used when analyzing the free-response answers provided by the participants were developed from the responses themselves using grounded theory (Martin & Tuner, 1986; Maxwell, 2013). First, categories of ideas emerged from the data for each survey item, and then codes were applied to the ideas. The data were then examined again and the codes were
compared for frequency of use. Common themes that emerged were then added as multiple choice response items for the particular item in the survey. This procedure was repeated for every item with free response answer options in the survey, and for each of the three survey versions. An example of a coding scheme from dissemination of the second version of the survey is presented in Appendix E.

Items with multiple choice responses were analyzed quantitatively to determine whether any answer options were redundant or occurred with a very low frequency. These answer options were investigated to determine whether there was an issue with language (students misinterpreting the meaning of the response) or significance (a phenomenon which did not exist within this population of students).

Human subjects

Institutional Review Board (IRB) approval was given prior to interviews and dissemination of the surveys. The approval letter is provided in Appendix A. Participants were given a consent form to sign denoting their agreement to have their data included in this study. No names were connected with the interview data, and any names provided in sample responses were pseudonyms. Students’ course grades were not influenced by participation in any part of this study, and instructors and teaching assistants did not know about any responses made by participants.

Data analysis and results

Development of the first version of the survey

The first version of the textbook use survey included a total of 11 items. The first item consisted of a consent statement for use of responses in the study. The 10 remaining items consisted of questions pertaining to the types of study materials used while studying, whether a
textbook was accessible to students, reasons for using a textbook to study, what activities were done while studying (e.g. taking notes or simply reading the text), and perceive instructor attitude toward the textbook. The first version of the textbook use survey can be found in Appendix F.

The 10 survey items were:
1. In the last week, I used the following resources when studying for chemistry:
2. The following describes the textbook I use most for studying:
3. The form of the textbook I use most for studying is:
4. The following best describes my access to the text:
5. The reason(s) I chose to use a textbook to study is/are:
6. Features of the textbook I used this week are (and how useful I found this feature):
7. When I used the textbook while studying I:
8. The total amount of time I spent studying for (or working on) chemistry this week:
9. When I study for chemistry I do it (and how often):
10. On a scale of 1-7, how much does my instructor value the textbook?

**Item 1**

The response options provided were based on study strategies of organic chemistry students reported by Lopez et al. (2013) as well as on comments made by the two interview participants. In their study, Lopez and coworkers found that students reported study strategies amongst a diverse group of students. Students used notes from class and handouts, read the course textbook, and read information from online sources. Item 1 focused on physical materials students would use to study rather than study activities, so self-reported activities such as speaking with a professor or peers were not included as answer options on this item of the survey.
Student B, when asked about what materials she uses while studying, reported only using class powerpoints and notes she wrote while in class. Student A also used powerpoints provided by her instructor, but she mentioned that she spent most time reviewing concepts and practicing problems using online homework. The final free response option was included to allow students to add to the list of possible study materials.

The first item contained a list of resources chemistry students use while studying for their chemistry courses, including the textbook, class slides, notes provided by the instructor, notes taken while in class, internet sources, practice problems from the textbook, online homework software, or something else (students were given the option to list up to three additional study materials as a free response answer). In addition, students were asked to estimate the percentage of study time they spent with each particular resource. This aspect of the item was included as a direct response to Student A’s comment that, though she used a number of materials while studying, she a vast majority of her time was spent with online homework software.

Items 2 and 3

In the syllabi for the general chemistry courses, the instructors suggested students purchase the 6th edition of Chemistry: The Molecular Nature of Matter and Change textbook by Silberberg. However, the instructors did not require a textbook to be purchased from the course, and encouraged use of an online textbook that was included with the purchase of online homework software. Based on this information, Item 2 asked students to describe the textbook they used to study, and two options were provided: “I use the textbook recommended by the professor,” or “I use a different textbook than the one recommended by the professor.” Item 3
then asked which textbook format the student used, either a physical textbook or an online version of a textbook. No free response answer options were available for these two items.

Item 4
Because textbook costs can be prohibitive for students to purchase them, there has been a growing interest in online or open educational resources for students (Smith & Jacobs, 2003; Robinson et al., 2014; Carnns, 2014). Item 4 addressed how students access the textbook they used while studying. The answer options to the question “Which of the following best describes my access to the text” were:

- I bought the text
- I am borrowing the text
- I am renting the text
- I use a text from the library
- I am sharing the text with another student
- I do not have access to a textbook because I do not feel I need a textbook for the course
- I do not have access to a textbook because I did not want to purchase a textbook
- Other (please specify) *this was a free-response option

The option of “sharing the text with another student,” was included in the possible answer options based on Student B’s response to prompt 4. She commented that she and a friend went in together to buy a textbook to share.
Item 5

Depending on the difficulty of the material or on the time of the semester, students may have differing motivations for turning to a textbook to study (Smith & Jacobs, 2003; Randahl, 2012; Lopez et al., 2013). Item 5 was designed to address external motivations for students to use textbooks while studying (as opposed to internal motivations such as self-regulation and self-efficacy). All answer options for this item were provided by comments made during interviews with Student A and Student B. The question item was “The reason(s) I chose to use a textbook to study is/are:” and students were allowed to choose multiple answers. The answer options were:

- to prepare for an upcoming exam
- to prepare for an upcoming quiz
- to review what I have learned in class
- to get a more thorough look at the material than what was presented in class
- to review what was not covered in class
- the instructor suggested it
- to review homework problems
- to help answer homework problems
- to practice problems
- other (please specify) *this was a free-response option

Student A, in response to prompt 1, reported reading the textbook regularly. She would read the textbook the evening before lecture class to review material from the previous lecture and to review material that would be covered in the next lecture. Student B reported not using the textbook much while studying, but that she would turn to example and practice problems in the text if she was studying for an exam. Each of the Students were asked to think about reasons why other people might use the textbook. Student A mentioned that if her professor suggested reading a certain section of the book, she would likely do that. Student B commented that some of her
friends would have the textbook open while working on online homework problems in case a worked out example in the textbook would help them answer a question online.

**Item 6**

Prior studies have focused on evaluating different textbook features (Gillespie, 1997; Russo, 1998; Smith & Jacobs, 2003; Taşdelen & Köseoğlu, 2008; Dávila & Talanquer, 2010; Pyburn & Pazicni, 2014). Item 6 was included to investigate what sections or features of the textbook were used when studied, and how useful that feature was considered by the student. The question item was: “Features of the physical or online textbook I used this week (select all that apply):”
The features listed as options were:

<table>
<thead>
<tr>
<th>Feature</th>
<th>Non-worked out follow-up problems in the chapter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction text</td>
<td></td>
</tr>
<tr>
<td>Summary text</td>
<td>End of chapter problems</td>
</tr>
<tr>
<td>Overview text (bulleted list of main points)</td>
<td>Solutions manual</td>
</tr>
<tr>
<td>Bolded words</td>
<td>Images or diagrams illustrating the chemical principles</td>
</tr>
<tr>
<td>Mathematical equations</td>
<td>Graphs or charts</td>
</tr>
<tr>
<td>Chemical equations</td>
<td>Other (please specify) *there were three opportunities for students to provide other features they used</td>
</tr>
<tr>
<td>Worked out example problems in the chapter</td>
<td></td>
</tr>
</tbody>
</table>

The option of solutions manual was included because Student B commented on using the solutions manual to help check her work when she would practice end-of-chapter problems.

Students were asked to indicate how useful they considered each textbook feature on a 1-7 Likert scale. Also, there was an option for each textbook feature for a student to write comments about why they did or did not find that feature to be useful.

**Item 7**

In order to elucidate what active study strategies students employed item 7 was included. Item 7 asked “When I used the textbook while studying this week, I (select all that apply);” The response options that were included were responses from the study by Lopez and coworkers about self-regulated learning study strategies (2013). These options included:
• wrote notes on my own paper
• wrote notes in the textbook
• highlighted the text in the book
• printed out online textbook pages
• worked out practice problems on my own paper
• just read, did not take notes
• did problems at the end of the chapter
• looked at problems, but did not work them out on paper

Students were also given the opportunity to report any other strategies they used while studying in a free-response section.

Items 8 and 9
Researchers may be interested in the time students take each week to set aside for studying, and what the study environment is like. Items 8 and 9 were included to investigate these potential questions.

Item 8 asked students to estimate the amount of time they had spent studying for chemistry during the past week. Options were:

Fewer than 3 hours  3-5 hours  5-7 hours  7-9 hours  9-11 hours  >11 hours

Item 9 asked “When I study for chemistry I do it.” and students were prompted to indicate whether they studied alone or in groups, and how often they studied in each situation. The answer options were

• alone
• in a group of 2-3 people
• in a group of more than 3 people
• only in PLTL groups
other (specify) *this option allowed for free-response answers.

For each of the options, students were asked to indicate whether they studied in those groups sometime, often, or most of the time. In the institution where this study took place, general chemistry courses have the option for students to participate in Peer Led Team Learning study groups (PLTL). The option of PLTL groups as an answer choice in Item 9 was added because Student B commented that she would review class material mainly with her PLTL group, which met once a week.

Item 10

Item 10 was included because of a response Student A had for prompt #4 during the interview. When asked whether she had access to a textbook, she responded that, although her instructor included a textbook in the syllabus in her first semester general chemistry course, the book was rarely referred to during class, and homework was never assigned from the book. Student A was asked why she thought the instructor did not refer to the textbook, and she responded that the instructor had commented that the textbook “wasn’t very good.” Based on these comments, a question about instructor attitude toward the textbook was included in the survey.

Results from dissemination of the first version of the survey and development of the second version

Based on the responses on the first version of the textbook use survey, alterations were made to Items 1-7. In addition, two items (2.11 and 2.12) were added to the survey. The second version of the survey is presented in APPENDIX G.

Item 1
Students were asked to approximate the amount of time spent with each study resource in terms of percentages in the first version of the survey. The responses showed that students would mark similar percentage of time amounts for all their choices. Therefore, the decision was made to change the “amount of time” choices from percentage of time to a Likert scale from 1 (Never) to 5 (Always).

**Items 2, 3, and 7**

Item 2 gave only two options for the question whether the student uses a textbook the instructor recommends or a different textbook. In this sample, 91% of students reported using the textbook recommended by the instructor, while the remaining 9% reported using a different textbook. Item 3 asked whether students used an electronic or physical version of the textbook. 73% of students reported using an electronic version of a textbook, while 27% reported using a physical textbook. Item 7 asked students what activities they engage in while studying, and included a range of possible options.

Based on results from Item 5 (“The reason I chose to use a textbook to study is”) it became clear that many students did not use a textbook while studying. Item 5 included a free-response option, and 24 out of 244 students (10%) reported not using a textbook in their free-response answer. Because of the responses to Item 5, the option of “I do not use a textbook while studying” was added to the answer choices for Items 2, 3, and 7. Because both Items 2 and 3 in the first version were forced-answer questions, students who did not use a textbook were required to report that they did use some sort of textbook. The results, then, were not indicative of the textbook usage for students in this sample.

**Item 5**
As mentioned above, Item 5 allowed for free-response answers to the prompt “The reason(s) I chose to use a textbook to study is/are (select all that apply).” In addition to students reporting not using a textbook at all, five students commented that they used the textbook to review material before the next lecture. Based on these responses, the options “To review before class” and “I did not use a textbook to study” were added to the answer choices.

Item 6

Item 6 asked students to identify what features found within the textbook they used, and which they viewed as being useful for learning. The free-response answers provided by students indicated that some alterations to the options listed should be made. Three students commented that they paid particular attention to words that were italicized in the text, so the answer option “Bolded words” was edited to read “Bolded or italicized words.” Two students reported reading all text sections and looking at all figures, and so the option “All text/sections in the chapter” was added. Finally, the option “Historical vignettes” was added based on a suggestion by a chemistry education colleague. Several students commented that they did not use any features of the textbook in the free-response section for this item, but no choice option for “I did not use a textbook” was added because this is not a forced-response item. Students who do not use any features of the textbook can simply not submit an answer.

New items

2.11

Two new items were added to the second version of the survey. The first new item was added to the survey between items 2 and 3, and will be referred to as Item 2.11. This new item asked students to comment on their reason(s) for choosing to not use a textbook as a study
material. This was a forced free-response item. Because many students reporting not using a textbook in their responses to the first version of the survey, the reasons for not using the textbook were worth exploring in the second version of the survey. The results from these responses would be coded into answer options for the third version of the survey.

2.12

The second item added to the survey was between items 7 and 8, and will be referred to as Item 2.12. This item was also a forced free-response item that asked students to comment on their experiences with and opinions about the role of textbooks in a learning environment. Item 2.12 was added as a response to several comments on Item 6 about the usefulness of features of the textbook. The phrases “not useful,” “boring,” “not relevant,” and “waste of time” were present in several comments. The addition of a question asking about opinions of the usefulness of textbooks in science courses was meant to probe why students would choose not to use a textbook to study.

Results from dissemination of the second version of the survey and development of the third version

Items 2, 3, 5, 6, 7, 8, 9, and 10 were not changed based on responses from the second version of the survey. Answers for Items 2.11 and 2.12 remained free-response in order to gather more data before developing the multiple choice answer options. Answer options were added to Items 1 and 4, and the wording of the question was altered for Item 1. A new item, Item 3.13, was added to the third version of the survey.

Item 1

When asked what resources students used while studying, the response “PLTL” occurred 15 times. The goal of the question was to determine study materials that students used rather than
resources such as peers, tutors, or instructors. The language used in the question was unclear based on student responses, and so the Item was reworded to ask, “In the last week I used the following materials when studying for chemistry, and used the resource approximately ________(select all that apply):” The options for the amount of time students spent with each material were on a Likert scale from 1 (Never) to 5 (Always).

A study material was also added to the answer options. Several mentions of a student study guide were made in the responses from the second version of the survey, and so “Student study guide” was added to the third version of the survey.

Item 4

Responses to Item 2.11 (“Please give your reason(s) for choosing not to use a textbook as a study material) informed another answer option that should be available for Item 4 (“Which of the following best describes my access to the text”). A large number of students (72 out of 209) made comments about not using the book they had purchased because it was not as helpful as other materials. These responses led to the answer option “I have access to a textbook, but choose not to use it while studying.”

Twenty five students responded that they did buy a textbook, but it was either used or was an e-book. The answer option “I bought the text” was then expanded into three questions for the third version: “I bought a new textbook,” “I bought a used textbook,” and “I bought the text along with online homework software.”

In the second version of the survey there was only one answer option for students who did not have access to the textbook, “I do not have access to the textbook because I do not feel I need a textbook for the course.” Based on the 15 student responses referring to the cost of the
textbook being a deciding factor for not having one, the option “I do not have access to a textbook because I did not want to purchase a textbook due to high cost” was added to Item 4.

New Item

Item 3.13

To get at student attitudes about the usefulness of the textbook, a final item was added at the end of the survey. Many of the free-response answers relayed negative feelings about textbooks. Common comments were that students can succeed without a textbook, textbooks are a waste of money, it was never used in class by the instructor, notes are all that are necessary to be successful in general chemistry, online textbooks were difficult to navigate, and that other resources are better than textbooks. There were a number of positive responses, including that textbooks help while studying for exams, they are good for practicing problems, and that they are useful for lab reports.

Item 3.13 was added to get an overall idea about a students’ beliefs about whether a textbook is useful for general chemistry. The item question was “Would you recommend a student coming into general chemistry next semester use a textbook as a study resource?” Students had to respond “yes” or “no,” and then provide a free-response explanation for their answer. The third version of the survey is presented in Appendix H.

Results from dissemination of the third version of the survey and development of the final version

Item 3.13

Item 3.13 asked students whether they would recommend future students get a textbook for general chemistry. 62% of students responded “no” and 38% responded “yes.” The students were then asked to give a reason for their answer. The answers were coded to generate the
answer options. Students could select more than one option that was relevant. The resulting answer options were:

- It couldn’t hurt.
- It can help explain ideas.
- It is useful for practice problems.
- I might have done better if I had used it.
- I didn’t use it and still did well in the course.
- It is a waste of money.
- The professor never used it.
- The professor’s notes are all that is necessary.
- It is not helpful.
- Other resources are more helpful.
- Tests are based on lecture, not the textbook.

**New Item, F.14**

Based on responses from item 2.12 (“Please share comments you may have on your experiences with, opinions on, or suggestions regarding textbooks and their role in a science learning environment”) Item F.14 was developed. Many students expressed the opinion that textbooks were not useful in their general chemistry experience. After coding, clusters of
explanations emerged. These clusters were translated into answer options for the Item “I choose not to use a textbook because (select all that apply).” Answer options are:

- My instructor does not use the textbook.
- My instructor does a better job explaining concepts than the book.
- Other sources were all that were necessary.
- It was too expensive.
- I found it to not be useful in previous courses.
- The online textbook was difficult to navigate.
- Textbooks are useful in some courses, but not chemistry.

The final version of the survey (APPENDIX I) included 12 items in the following order:

1. (Item 1) In the last week, I used the following materials when studying for chemistry, and used that resource how often:
2. (Item 2) The following describes the textbook I use most for studying:
3. (Item F.14) If you do not use a textbook for studying, please answer the following question. I choose not to use a textbook because (select all that apply).
4. (Item 3) The form of the textbook I use most for studying is:
5. (Item 4) Which of the following best describes my access to the text?
6. (Item 5) The reason(s) I chose to use a textbook to study is/are (select all that apply):
7. (Item 6) Features of the textbook I used this week are (and how useful I found this feature):
8. (Item 7) When I used the textbook while studying I (select all that apply):
9. (Item 8) The total amount of time I spent studying for (or working on) chemistry this week:
10. (Item 9) When I study for chemistry I do it (and how often):
11. (Item 10) On a scale of 1-7, how much does my instructor value the textbook?
12. (Item 3.13) Would you recommend a student coming into general chemistry next semester use a textbook as a study resource? Why?
Discussion and conclusions

The final textbook use survey was developed after three iterations of surveys were constructed and disseminated. Based on student responses of each iteration, items and answer options were reworded, omitted, and added to reflect how students interpreted each item question. Student-led development of a survey allows for the survey and answer options to be better understood by students. This study aimed to use terminology and ideas credible to general chemistry students, and it achieved that by using student-generated explanations and language, leading to trustworthiness of the survey items in terms of credibility (Guba, 1981). In qualitative research, credibility is considered to be analogous to internal validity (Maxwell, 2013).

The final version of the textbook survey was meant to be comprehensive, brief, and accurate. The items were added to the survey when free response answers hinted at uses for textbooks or reasons for not using textbooks that previous survey versions had missed. The survey achieved its goal of being brief by omitting redundant items (responses from Items 2.12 and 2.11 were incorporated into other items), and resulting in a survey with 12 total items. Accuracy of the final version cannot be addressed at this stage. Further work must be done to address the correlation of student self-reported responses on the survey with their actual study habits. A future study can include interviews with students to verify the accuracy of their self-reports.

Though not the purpose of this study, there were some interesting findings regarding student use of and attitudes about textbooks when the responses from each iteration of the survey were analyzed. At the beginning of the semester, 73% of students reported using an online textbook while 27% reported using a physical textbook. However, when the option of “I did not use a textbook” was included in the second version, 48% of students chose that response, while
29% of students used an online textbook and 22% used a physical textbook. Later in the semester, one week before the final exam, the percentage of students not using a textbook rose to 60%. The finding that students are less likely to use textbooks later in the semester was also observed in a study by Bunce et al. (2017).

When asked about access to a textbook, 51% of participants reported purchasing a textbook (18% bought a new textbook, 20% purchased the online textbook that was included with the electronic homework, and 13% purchased a used textbook). In addition, 13% of respondents reported having bought a textbook, but not using it during the semester. This raises the question of why students buy the expensive textbook without using it in any capacity. It is possible that this 13% of students purchased the required online homework software which included access to a textbook, in which case the money was not “wasted.”

Materials students used most often were class notes, notes provided by the professor, and online homework, while the textbook, problems from the textbook, and the student study guide were used rarely or never, as illustrated in Figure 3.2. This finding has implications for the instructors of this population of general chemistry students: Although online and physical textbook resources were available to students, the majority of students did not use these materials at all, and if the textbook materials were used for studying, they were used very rarely. If the instructors believe that using a textbook is useful for students, they may consider assigning work or readings from the textbook. Otherwise, students will refer to notes written during the semester.

The Textbook Use Survey contains items that can address a number of dimensions regarding textbook use. Student preferences between online and physical textbook options can emerge through the form of the textbook students tend to use. Different students may use
different textbook features. An interesting research question could be whether there is a relationship between course performance or conceptual understanding and the types of study materials or activities an individual chooses to use. Another relationship to investigate could be between student achievement and attitudes toward the textbook (Items 5 and 3.13). This survey is meant to be a resource for quantitatively approaching research questions involving student study habits involving the textbook.

Figure 3.2 A histogram displaying frequency of use of different materials for studying.

**Implications**

This survey was developed to be used by instructors or researchers to help gather quantitative information from students about use and attitudes about textbooks. Currently there are no surveys available for widespread dissemination about textbook use, and many studies collect qualitative data from interviews, which is very time intensive. The quick survey developed here can allow for less time consuming data collection. For researchers it means less time spent coding interview transcripts. For instructors, it could mean that they could revise how
to discuss studying with their class if the instructor notices that certain valuable materials are not being utilized by students. Or if a relationship is found between achievement and study habits using the textbook, the instructor can relay that information to students so that students may choose to try different strategies.

**Limitations**

This study relied on student self-reported study habits, which may not be accurate. Student responses were not anonymous, as students received a small amount of course credit for completing the survey. In each dissemination of the survey there were participants who merely answered the first item in order to earn credit for completion. Other survey responses were clearly made with no effort. For example, in the third iteration of the survey the item “Please give your reason for choosing not to use a textbook as a study material,” 15 students responded simply with “No.” These students may not have put much thought into the rest of their responses, but there is no way for the researcher to know which answers were genuine and which were simply given for the appearance that the survey was completed.

Another limitation is that interviews were not conducted after the final survey was developed. Interviews would give insight into student interpretation of the final items and answer options and would increase the validity that the survey is measuring what is intended.

The biggest limitation is that the survey was developed using responses from a single sample of students at one university in the Northeast. The group of participants did not represent a very diverse range of backgrounds, and so experiences with textbooks may not vary widely. Also, the environment of a chemistry classroom is affected by instructor, university culture, and geographic location. The textbook use survey should be disseminated to a variety of populations.
to determine whether the answer options available encompass a variety of students’ experiences with and attitudes about textbooks.
CHAPTER 4
SUMMARY AND FUTURE WORK

The results in Chapter 1 show that students with higher prior knowledge and reading comprehension abilities are predicted to perform better on post-tests about general chemistry topics, and that reading a text passage about those topics does not affect post-test scores. The interaction between reading a text and prior knowledge was shown to not be statistically significant at 95% confidence, but there is evidence that this interaction, referred to as an “expertise reversal,” is beginning to emerge as a possible factor in post-test performance. A limitation of this study is that the text passages used in the intervention had very high levels of referential cohesion, which has been shown in previous studies to preferentially aid students with low prior knowledge (Kendeo, et al., 2003). Students with higher prior knowledge benefit more from having to make connections between ideas independently, and text with high cohesion does not allow for that independent thought (O’Reilly & McNamara, 2007). Future work into the effects of text characteristics on learning gains may give more insight into the expertise reversal effect. Text passages containing the same information but presented with different levels of cohesion can be investigated to see whether an expertise reversal effect is present when reading to learn. However, it may be challenging to present the same information at differing levels of cohesion.

Future work may also include using text formats such as conceptual change texts (or refutation texts) (Tippett, 2010) to identify whether there really exists an interaction between
reading text and prior knowledge, and whether this interaction depends on the text type.
Conceptual change texts are written in a way where common misconceptions are explicitly pointed out, and then a discussion explains why the misconceptions do not accurately represent the concept. It would be hypothesized that these texts, written with different degrees of cohesion, may also help with learning gains as students would be confronted with the deficiencies of their own conceptions, which is the first step in the conceptual change framework (Posner et al., 1982). On way to provide various types of text based on student prior knowledge can be achieved through online platforms, where the student can take a short assessment to determine level of knowledge, and then be led to a particular text passage.

Chapter 2 investigated the epistemological beliefs of students in a first-semester college general chemistry course, most of whom were first-year college students. Overall, students were found to hold beliefs that fall between “Transitional” and “Independent” understanding of epistemology, based on Baxter Magolda’s levels of conceptualization of epistemology (2004). It was found that course performance (as measured by midterm or final exam scores) was a statistically significant predictor of total EBAPS scores at both time points tested in this study (B = 2.45, p < 0.001 for the beginning of the semester; B = 2.97, p = 0.017 for the end of the semester). Students who perform higher on these assessments also scored slightly higher on the epistemological beliefs assessment. Only a small number of students held beliefs at the most “naïve” level of “Absolute Knowing,” where it is understood that learning only happens when an expert imparts knowledge on the learner and that scientific concepts are concrete and absolute. However, there is room for epistemological belief growth. This study was limited in scope, as only class performance, sex, and lecturer were compared to epistemological beliefs, and only students in a general chemistry course were assessed. Future work could investigate how
epistemological beliefs develop over time at university and graduate school. Also, students have
the opportunity to be involved in many academic activities, and the effects of these activities on
epistemological belief growth would be particularly interesting. If it is found that some activities
correlate with epistemological belief growth, instructors can encourage participation in those
activities, or provide those opportunities to students.

A student’s prior knowledge, reading skill, and epistemological beliefs may influence
how they choose to study, and what study materials they use. Chapter 3 discussed the
development of a survey that instructors and researchers can use to identify how and why
students use textbook materials when studying general chemistry concepts. The development of
textbook materials, both physical and online, should take into account student preferences and
habits. Future work could include a large scale study at different universities (with different
population demographics). Comparisons in textbook use, instructor attitudes towards textbooks,
and preferences over textbook features would give valuable information for textbook developers
and instructors. Also, if it is found that students are not interacting with features of the textbook
that are particularly informational (or deemed important by the instructor), instructors can know
to guide their students to these features. A future study that identifies where students gaze when
reading an online textbook would give greater insight into what the student is doing while
studying in real time. Such an eye-tracking study, which would compare student gaze with
performance on an assessment about the topic they are reading about, could provide information
about design of text materials.


Cumming, G. & Calin-Jageman, R. ESCI for meta-analysis, available online https://thenewstatistics.com/itns/esci/


Epistemological Beliefs Assessment for Physical Science (EBAPS). Online access: http://www2.physics.umd.edu/~elby/EBAPS/home.htm


APPENDIX A

University of New Hampshire
Research Integrity Services, Service Building
51 College Road, Durham, NH 03824-3585
Fax: 603-862-3564

05-Sep-2014

Buell, Rene
Chemistry, Parsons Hall
1102 Bennett Way
Newmarket, NH 03857

IRB #: 6072
Study: Investigation of the Effects of Student Reading Level and Text Type on Conceptual Change Activation
Approval Date: 04-Sep-2014

The Institutional Review Board for the Protection of Human Subjects in Research (IRB) has reviewed and approved the protocol for your study as Exempt as described in Title 45, Code of Federal Regulations (CFR), Part 46, Subsection 101(b). Approval is granted to conduct your study as described in your protocol.

Researchers who conduct studies involving human subjects have responsibilities as outlined in the attached document, Responsibilities of Directors of Research Studies Involving Human Subjects. (This document is also available at http://unh.edu/research/irb-application-resources.) Please read this document carefully before commencing your work involving human subjects.

Upon completion of your study, please complete the enclosed Exempt Study Final Report form and return it to this office along with a report of your findings.

If you have questions or concerns about your study or this approval, please feel free to contact me at 603-862-2003 or Julie.simpson@unh.edu. Please refer to the IRB # above in all correspondence related to this study. The IRB wishes you success with your research.

For the IRB,

Julie F. Simpson
Director

cc: File
Pazicni, Samuel
20-Aug-2014

Buell, Rene
Chemistry, Parsons Hall
1102 Bennett Way
Newmarket, NH 03857

IRB #: 6053
Study: Investigation of Student Textbook Usage in General Chemistry
Approval Date: 20-Aug-2014

The Institutional Review Board for the Protection of Human Subjects in Research (IRB) has reviewed and approved the protocol for your study as Exempt as described in Title 45, Code of Federal Regulations (CFR), Part 46, Subsection 101(b). Approval is granted to conduct your study as described in your protocol.

Researchers who conduct studies involving human subjects have responsibilities as outlined in the attached document, Responsibilities of Directors of Research Studies Involving Human Subjects. (This document is also available at http://unh.edu/research/irb-application-resources.) Please read this document carefully before commencing your work involving human subjects.

Upon completion of your study, please complete the enclosed Exempt Study Final Report form and return it to this office along with a report of your findings.

If you have questions or concerns about your study or this approval, please feel free to contact me at 603-862-2003 or Julie.simpson@unh.edu. Please refer to the IRB # above in all correspondence related to this study. The IRB wishes you success with your research.

For the IRB,

Julie F. Simpson
Director

cc: File
Pazinčí, Samuel
In order to make compounds, atoms and ions come together to form chemical bonds. The components of a chemical bond (i.e. atoms or ions) interact via electrostatic attractions, with positively charged particles attracted to negatively charged particles. The bonds of most compounds fall into one of two categories: ionic and covalent. It must be noted, however, that purely covalent and purely ionic interactions are two extreme cases of chemical bonding, and in reality all compounds exhibit features of both types of bonding to some degree.

As suggested by the name, ions form ionic bonds. In ionic species, oppositely charged ions are attracted to each other and form a solid three dimensional array (or lattice) of particles. Metals, with low ionization energies, have a tendency to lose electrons; nonmetals, with high electron affinities, have a tendency to gain electrons. Once a metal has ionized, it becomes a positively charged cation, while a nonmetal becomes a negatively charged anion. When a collection of oppositely-charged ions are close to each other, they will be attracted by electrostatic forces, and ionic bonds will form. The resulting solid is made up of an ordered arrangement of ions interacting with one another—each cation is attracted to several neighboring anions, and each anion interacts with several neighboring cations. The interaction between ions in an array is not limited to just monoatomic ions. Ionic bonds can also be generalized to include ions that are made up of several atoms, or polyatomic ions.

Unlike ionic bonding cases, none of the atoms involved in a covalent bond have gained or lost electrons. Instead, covalent compounds are composed of neutral atomic species and not ions. A covalent bond can be described as a “sharing” of electrons between the nuclei of two atoms, and compounds composed of covalent bonds are commonly called “molecules”. Molecules are discreet networks of covalently bonded atoms that range in size according to how many atoms comprise the compound; only in very rare cases do covalent compounds contain enough atoms to rival the infinite lattice structures of ionic compounds.

Though electrons involved in a covalent bond are localized between two atoms, the extent to which the electrons are “shared” varies. That is, one of the atoms in the bond may attract the electrons more to itself because the atom possesses a greater electronegativity relative to the other atom involved in the bond. So, electrons between atoms of unequal electronegativity may in fact lie closer to the more electronegative atom. In this case, the electrons are still shared between the two nuclei, but in an unequal manner; this scenario is known as a polar covalent bond. If the covalently bonded atoms have the same electronegativity value, the electrons between them will be shared equally.
Redox concepts text passage

Many chemical reactions can be categorized as one of three major classes: Precipitation, acid-base, or oxidation-reduction (redox) reactions. This passage will focus on the latter class.

Redox reactions involve the transfer of electrons between chemical species, and these processes can occur among ionic or covalently bonded compounds as well as pure elemental species. Electron transfer occurs in many important applications, including batteries, combustion, photosynthesis, electroplating, and cellular respiration.

In order to identify a reaction as a redox process, electrons must move from one reagent to another. A bookkeeping method has been devised by chemists to keep track of whether an electron has been “gained” or “lost” by an atom in a reaction by assigning oxidation numbers (or oxidation states) to individual atoms. This is not to be confused with assignment of charge to an atom or molecule. The oxidation state is not a true physical charge which can be measured, but instead is the theoretical charge the atom would have if the atom was ionically bonded with the other atoms in the species. This scheme is followed for molecules exhibiting the whole spectrum of bonding, from purely ionic to purely covalent, and is used simply for the ease of keeping track of electrons.

Atoms in their pure elemental form have an oxidation number of zero, but if the atoms are charged ions or are bound to other atoms, then a positive or negative oxidation number can be assigned. The oxidation number of an atom in its ionic form will be the same as the charge the ion carries. If the atom is a group 1 metal, it will have an oxidation number of +1, and atoms in group 2 will have an oxidation number of +2. When the atom is a halogen the oxidation number for that atom will be -1. Oxygen will usually have an oxidation number of -2, while hydrogen typically is assigned to an oxidation number of +1. The sum of all the oxidation numbers assigned to atoms in a compound must be equal to the overall charge of the compound. For example, in phosphate (PO₄³⁻) the oxygen atoms are each assigned to a -2 O.N. This will make the phosphorous atom have a +5 oxidation number so that overall the phosphate ion will have a 3- charge.

When a redox process occurs, the oxidation states for at least two atoms will change during the course of the reaction. If the oxidation state of an atom increases to become more positive (or less negative) after the reaction, then the species containing that atom will have been oxidized (or there will have been a loss of electrons from that species). If the oxidation state of the atom becomes less positive (or more negative), then that species has been reduced, or has gained electrons during the reaction. Commonly oxidation is referred to as a loss of electrons, while reduction is referred to as a gain of electrons. The species which undergoes oxidation is called the “reductant” or “reducing agent,” and the species which is reduced is called the “oxidant” or “oxidizing agent. The reductant and oxidant work together in a redox reaction.

The chemical equation must be balanced to determine the number of electrons transferred between the reductant and oxidant. Balancing a redox reaction involves splitting the chemical equation into two half-reactions, where one half-reaction describes the oxidation process, and the other half-reaction describes the reduction process. The electrons transferred in each half reaction are found independently, and then the total number of electrons involved in the overall redox process is the common multiple of the number of electrons involved in the separate half-reactions.
APPENDIX D

Consent form for participating in textbook use study

Dear Chem xxx student:

I am inviting you to participate in a research project to investigate how general chemistry students use their textbooks while studying for the course. Chemistry instructors can use this information to learn more about their students and design aspects of their course to help students with study habits. By gaining more information about student characteristics, instructors will be able to better tailor the design of their course to their student population. I plan to work with approximately 600 students per semester in this study. You must be at least 18 years old to participate in this study.

If you agree to participate in this study, you will be agreeing to allow us to use your normal UNH academic information and chemistry course grades and work products, in particular your:

- Chem 403 or Chem 405 course exam scores in class and group averages
- Regular course work (assignments, laboratory reports, quizzes, and exams) to assess understanding
- Your responses to any surveys we provide
- Your student ID number (strictly as a means for tracking and linking information; always confidential)
- Entrance exam scores (SAT, ACT, UNH mathematics placement exam)

You will be asked to complete several short surveys online over the course of the semester about study habits and textbook usage. Completion of each textbook usage survey should take approximately 20 minutes. Completion of all surveys is required as an assignment for the course. You will receive credit towards an assignment as long as you complete the surveys. Although completing the surveys is required, participating in the study is optional. By participating, you are allowing the researchers to use your responses to the surveys.

As a follow up some of you may be selected for a brief interview. In order to have an accurate record of your responses, we ask to collect an audio record for the interview. In any presentation of information we collect from you, you will not be identified by name. To protect your identity, records of your interview will be given an anonymous code. Recordings will be saved and kept on secure computers for future research.

One risk involved the remote chance that your work might be identified by name. To guard against this, statistical analyses will be carried out on group averages; no individuals will be identified. Code names will be used when presenting examples of student work.

A benefit for participating in this study includes being given an opportunity to reflect on your study skills. Ultimately, future students will benefit from this study because instructors will be able to use the results of the textbook usage survey to learn more about their students, and will therefore have this information to make decisions on how they design their course.
Participation in this study is strictly voluntary; your refusal to participate will involve no prejudice, penalty, or loss of benefits to which you would otherwise be entitled. If you agree to participate you may withdraw at any time during the study. However, if you do not participate in the study or complete the survey you will not be entered into the raffle.

I seek to maintain the confidentiality of all data and records associated with your participation in this research. There are, however, rare instances when I am required to share personally-identifiable information (e.g., according to policy, contract, regulation). For example, in response to a complaint about the research, officials at the University of New Hampshire, designees of the sponsor(s), and/or regulatory and oversight government agencies may access research data. Further, any communication via the Internet poses minimal risk of a breach of confidentiality. Data will be stored on a computer and only Dr. Samuel Pazicni, your course instructor, and I will have access to this data. I will report the data in aggregate. The results will be used in reports, presentations, and publications.

If you have any questions about this research project or would like more information before, during, or after the study, you may contact Samuel Pazicni at 603-862-2529. If you have questions about your rights as a research subject, you may contact Dr. Julie Simpson in UNH Research Integrity Services at 603-862-2003 or Julie.simpson@unh.edu to discuss them.

PLEASE PROVIDE YOUR FULL NAME IN THE TEXT BOX BELOW. CLICK THE "SUBMIT" BUTTON. WHEN THE NEW PAGE LOADS, INDICATE IF YOU GIVE PERMISSION FOR YOUR SCORES ON THE CHEMISTRY COMPREHENSION BATTERY TO BE INCLUDED IN THIS STUDY. CLICK "SUBMIT" TO CONFIRM YOUR RESPONSE.

Sincerely,

René Buell
Ph.D. Student
Department of Chemistry
## APPENDIX E

Example coding scheme

<table>
<thead>
<tr>
<th>Response to “Please give your reason for choosing not to use a textbook as a study material.”</th>
<th>Main ideas present in comment</th>
<th>Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>I do not learn from math or chemistry textbooks easily. I learn better for hearing someone explain it to me.</td>
<td>-Learn better from hearing than reading</td>
<td>Not my learning style</td>
</tr>
<tr>
<td>i feel as though the lecture gives me all the information i need to succeed</td>
<td>-Not necessary to succeed</td>
<td>Not necessary to succeed</td>
</tr>
<tr>
<td>-Lecture materials are enough</td>
<td>Instructor explains enough</td>
<td></td>
</tr>
<tr>
<td>I don't use textbook because the notes that i take in class are much more precise and to the point, and make studying more effective and efficient.</td>
<td>-Notes are enough</td>
<td>Not my learning style</td>
</tr>
<tr>
<td>-Reading book is not efficient</td>
<td>Other materials are better</td>
<td></td>
</tr>
<tr>
<td>I couldn't find it and afford it</td>
<td>-Too expensive</td>
<td>Cost</td>
</tr>
<tr>
<td>I do not feel it is necessary because all of the information we need for tests is covered in lecture. Tests are lecture based</td>
<td>-Lecture materials are enough</td>
<td>Instructor explains enough</td>
</tr>
<tr>
<td>-Tests not based on textbook</td>
<td>Not necessary to succeed</td>
<td></td>
</tr>
<tr>
<td>-Not necessary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>It's an online textbook, and it confuses me more than simply learning the materials in other ways.</td>
<td>-Online text is confusing</td>
<td>Not my learning style</td>
</tr>
<tr>
<td>-Other methods of learning are better</td>
<td>Don't like online text</td>
<td></td>
</tr>
<tr>
<td>I never bought it and I don't like using the ebook</td>
<td>-Don’t like online text</td>
<td>Don’t like online text</td>
</tr>
<tr>
<td>I am able to understand my notes better, and I did not purchase one.</td>
<td>-Notes are enough</td>
<td>Other materials are better</td>
</tr>
<tr>
<td>I use the power points</td>
<td>-Use instructor notes</td>
<td>Instructor explains enough</td>
</tr>
<tr>
<td>I don't need it</td>
<td>-Not necessary</td>
<td>Not necessary to succeed</td>
</tr>
<tr>
<td>the subscription to the online textbook was from last semester when the professor taught with it, but I dont use it this semester because professor said it is not needed and everything on the exam will come from lecture notes</td>
<td>-Professor did not use it</td>
<td>Instructor explains enough</td>
</tr>
<tr>
<td>-Notes are enough</td>
<td>Not necessary to succeed</td>
<td></td>
</tr>
<tr>
<td>-Exam is based on lecture</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX F
Textbook use survey, version 1

Consent to use your responses in the study

☐ I DO give permission for my survey and interview responses to be used in the study.

☐ I DO NOT give permission for my survey and interview responses to be used in the study.

In the last week, I used the following resources when studying for chemistry, and spent approximately ___ percent of my time with that resource:

<table>
<thead>
<tr>
<th>I used this resource</th>
<th>Percent of study time spent with this resource</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10%  20%  30%  40%  50%  60%  70%  80%  90%  100%</td>
</tr>
</tbody>
</table>

---
<table>
<thead>
<tr>
<th>Resource</th>
<th>Circle 1</th>
<th>Circle 2</th>
<th>Circle 3</th>
<th>Circle 4</th>
<th>Circle 5</th>
<th>Circle 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Textbook</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C</td>
</tr>
<tr>
<td>Powerpoint slides from class</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C</td>
</tr>
<tr>
<td>Notes provided by instructor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C</td>
</tr>
<tr>
<td>Notes I wrote while in class</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C</td>
</tr>
<tr>
<td>Text about the topic from the internet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C</td>
</tr>
<tr>
<td>Homework problems from textbook</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C</td>
</tr>
<tr>
<td>Online homework software (ALEKS, Sapling, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C</td>
</tr>
<tr>
<td>Something else (please specify)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>C</td>
</tr>
</tbody>
</table>

The following best describes the textbook I use most for studying:

- [ ] I use the textbook recommended by the professor
- [ ] I use a different textbook than the one recommended by the professor
The form of the textbook I use most for studying is

- the physical textbook
- an online version of the textbook

For the remainder of the survey, answer with respect to the text which you primarily use while studying.

Which of the following best describes my access to the text?

- I bought the text
- I am borrowing the text
- I am renting the text
- I use a text from the library
- I am sharing the text with another student
- I do not have access to a textbook because I do not feel I need a textbook for the course
- Other (please specify) ________________________________
The reason(s) I chose to use a textbook to study is/are (select all that apply):

- To prepare for an upcoming exam
- To prepare for an upcoming quiz
- To review what I have learned in class
- To get a more thorough look at the material than what was presented in class
- To review what was not covered in class
- The instructor suggested it
- To review homework problems
- To help answer homework problems
- To practice problems
- Other (please specify) _____________________________

Features of the physical or online textbook I used this week (select all that apply):

<table>
<thead>
<tr>
<th>I used this feature</th>
<th>Not at all useful</th>
<th>2</th>
<th>3</th>
<th>Somewhat Useful</th>
<th>5</th>
<th>6</th>
<th>Extremely useful</th>
<th>Why I found this feature to be useful/not useful:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-worked out follow-up problems in the chapter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>End of chapter problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solutions manual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text with real-world applications or examples</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Images or diagrams illustrating the chemical principles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graphs or charts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td></td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other (please specify):</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
<td>o</td>
</tr>
</tbody>
</table>
When I used the textbook while studying this week, I (select all that apply):

☐ Wrote notes on my own paper

☐ Wrote notes in the textbook

☐ Highlighted the text in the book

☐ Printed out online textbook pages

☐ Worked out practice problems on my own paper

☐ Just read, did not take notes

☐ Did problems at the end of the chapter

☐ Looked at problems, but did not work them out on paper

☐ Other (please specify): ________________________________________________

☐ Other (please specify): ________________________________________________

☐ Other (please specify): ________________________________________________

☐ Other (please specify): ________________________________________________
The total amount of time I spent studying for (or working on) chemistry the past week was:

- [ ] Fewer than 3 hours
- [ ] 3-5 hours
- [ ] 5-7 hours
- [ ] 7-9 hours
- [ ] 9-11 hours
- [ ] More than 11 hours

When I study for chemistry I do it:

<table>
<thead>
<tr>
<th>Size of study group</th>
<th>Sometimes</th>
<th>Often</th>
<th>Most of the time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alone</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>In a group of 2-3 people</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>In a group with more than 3 people</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Only in PLTL groups</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Other (please specify):</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
</tbody>
</table>
On a scale of 1-7, how much does my instructor value the textbook?

- Does not value the textbook at all
- 2
- 3
- Somewhat values the textbook
- 5
- 6
- Very highly values the textbook
APPENDIX G
Textbook use survey, version 2

Consent to use your responses in the study

☐ I DO give permission for my survey and interview responses to be used in the study.

☐ I DO NOT give permission for my survey and interview responses to be used in the study.

In the last week, I used the following resources when studying for chemistry, and spent approximately ___ amount of time using that resource (select all that apply):

<table>
<thead>
<tr>
<th>How often I used this resource</th>
</tr>
</thead>
<tbody>
<tr>
<td>I used this resource while studying</td>
</tr>
</tbody>
</table>
The following best describes the textbook I use most for studying:

- I use the textbook recommended by the professor.
- I use a different textbook than the one recommended by the professor.
- I do not use a textbook while studying.
Display This Question:

If The following best describes the textbook I use most for studying: = I do not use a textbook while studying.

Please give your reason(s) for choosing not to use a textbook as a study material:

________________________________________________________________

The form of the textbook I use most for studying is

- [ ] the physical textbook
- [ ] an online version of the textbook
- [ ] I do not use a textbook while studying

For the remainder of the survey, answer with respect to the text that you primarily use while studying.

Which of the following best describes my access to the text?

- [ ] I bought the text
- [ ] I am borrowing the text
- [ ] I am renting the text
- [ ] I use a text from the library
- [ ] I am sharing the text with another student
- [ ] I do not have access to a textbook because I do not feel I need a textbook for the course
- [ ] I do not have access to a textbook because I did not want to purchase a textbook
- [ ] I have access to a textbook, but choose not to use it while studying
- [ ] Other (please specify) ________________________________
The reason(s) I chose to use a textbook to study is/are (select all that apply):

- To prepare for an upcoming exam
- To prepare for an upcoming quiz
- To review material before seeing it in class
- To review what I have learned in class
- To get a more thorough look at the material than was presented in class
- To review what was not covered in class
- The instructor suggested it
- To review homework problems
- To help answer homework problems
- To practice problems
- I do not use a textbook
- Other (please specify)  ________________________________________________

Features of the physical or online textbook I used this week (select all that apply):

<table>
<thead>
<tr>
<th>I used this feature</th>
<th>Not at all useful</th>
<th>Somewhat Useful</th>
<th>Extremely useful</th>
<th>Why I found this feature to be useful/not useful:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>All text/sections in the chapter</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Introduction text</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summary text</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overview text (bulleted list of main points)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text with real-world applications or examples</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical references or vignettes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bolded or italicized words</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>Grid 1</td>
<td>Grid 2</td>
<td>Grid 3</td>
<td>Grid 4</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>Mathematical equations</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Chemical equations</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Worked out example problems in the chapter</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Non-worked out follow-up problems in the chapter</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>End of chapter problems</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Images or diagrams illustrating the chemical principles</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Graphs or charts</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Solutions manual</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Other (please specify)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify):</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify):</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
When I used the textbook while studying this week, I (select all that apply):

- [ ] Wrote notes on my own paper
- [ ] Wrote notes in the textbook
- [ ] Highlighted the text in the book
- [ ] Printed out online textbook pages
- [ ] Worked out practice problems on my own paper
- [ ] Just read, did not take notes
- [ ] Did problems at the end of the chapter
- [ ] Looked at problems, but did not work them out on paper
- [ ] I did not use the textbook to study
- [ ] Other (please specify): ________________________________
- [ ] Other (please specify): ________________________________
- [ ] Other (please specify): ________________________________
- [ ] Other (please specify): ________________________________

Please share comments you may have on your experiences with, opinions on, or suggestions regarding textbooks and their role in a science learning environment.

______________________________________________________________________________________________________________
The total amount of time I spent studying for (or working on) chemistry the past week was:

- [ ] Fewer than 3 hours
- [ ] 3-5 hours
- [ ] 5-7 hours
- [ ] 7-9 hours
- [ ] 9-11 hours
- [ ] More than 11 hours

When I study for chemistry I do it:

<table>
<thead>
<tr>
<th>Size of study group</th>
<th>Sometimes</th>
<th>Often</th>
<th>Most of the time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alone</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>In a group of 2-3 people</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>In a group with more than 3 people</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Only in PLTL groups</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Other (please specify):</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
</tbody>
</table>
On a scale of 1-7, how much does my instructor value the textbook?

- Does not value the textbook at all
- 2
- 3
- Somewhat values the textbook
- 5
- 6
- Very highly values the textbook
APPENDIX H
Textbook use survey, version 3

Consent to use your responses in the study

☐ I DO give permission for my survey and interview responses to be used in the study.

☐ I DO NOT give permission for my survey and interview responses to be used in the study.

In the last week, I used the following materials when studying for chemistry, and used the resource approximately ______ (select all that apply):

<table>
<thead>
<tr>
<th>Click to write Column 1</th>
<th>How often I used this material</th>
</tr>
</thead>
<tbody>
<tr>
<td>I used this material while studying</td>
<td>Never</td>
</tr>
<tr>
<td>Resource</td>
<td>0</td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>----</td>
</tr>
<tr>
<td>Textbook</td>
<td>O</td>
</tr>
<tr>
<td>Student Study Guide</td>
<td>O</td>
</tr>
<tr>
<td>Powerpoint slides from class</td>
<td>O</td>
</tr>
<tr>
<td>Notes provided by the instructor</td>
<td>O</td>
</tr>
<tr>
<td>Notes I wrote while in class</td>
<td>O</td>
</tr>
<tr>
<td>Text about the topic from the internet</td>
<td>O</td>
</tr>
<tr>
<td>Homework problems from the textbook</td>
<td>O</td>
</tr>
<tr>
<td>Online homework software (ALEKS, Sapling, etc.)</td>
<td>O</td>
</tr>
<tr>
<td>Previous exams or quizzes</td>
<td>O</td>
</tr>
<tr>
<td>Something else (please specify)</td>
<td>O</td>
</tr>
</tbody>
</table>

The following best describes the textbook I use most for studying:

- [ ] I use the textbook recommended by the professor.
- [ ] I use a different textbook than the one recommended by the professor.
- [ ] I do not use a textbook while studying.
Display This Question:

If The following best describes the textbook I use most for studying: = I do not use a textbook while studying.

Please give your reason(s) for choosing not to use a textbook as a study material:

________________________________________________________________

The form of the textbook I use most for studying is

○ the physical textbook

○ an online version of the textbook

○ I do not use a textbook while studying

For the remainder of the survey, answer with respect to the text that you primarily use while studying.
Which of the following best describes my access to the text?

- I bought a new textbook
- I bought a used textbook
- I bought the text along with online homework software
- I am borrowing the text
- I am renting the text
- I use a text from the library
- I am sharing the text with another student
- I do not have access to a textbook because I do not feel I need a textbook for the course
- I do not have access to a textbook because I did not want to purchase a textbook
- I do not have access to a textbook because I did not want to purchase a textbook due to high cost
- I have access to a textbook, but choose not to use it while studying
- Other (please specify)  ________________________________________________
The reason(s) I chose to use a textbook to study is/are (select all that apply):

☐ To prepare for an upcoming exam
☐ To prepare for an upcoming quiz
☐ To review material before seeing it in class
☐ To review what I have learned in class
☐ To get a more thorough look at the material than was presented in class
☐ To review what was not covered in class
☐ The instructor suggested it
☐ To review homework problems
☐ To help answer homework problems
☐ To practice problems
☐ I do not use a textbook
☐ Other (please specify) ________________________________________________

Features of the physical or online textbook I used this week (select all that apply):

<table>
<thead>
<tr>
<th></th>
<th>How useful is this feature in helping me understand the concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>I used this feature</td>
<td>Not at all useful 2 3 Somewhat Useful 5 6 Extremely useful</td>
</tr>
</tbody>
</table>

Why I found this feature to be useful/not useful:
<table>
<thead>
<tr>
<th>Section Type</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>All text/sections in the chapter</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Introduction text</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Summary text</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Overview text (bulleted list of main points)</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Text with real-world applications or examples</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Historical references or vignettes</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Bolded or italicized words</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Mathematical equations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemical equations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Worked out example problems in the chapter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-worked out follow-up problems in the chapter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>End of chapter problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Images or diagrams illustrating the chemical principles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graphs or charts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solutions manual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Other (please specify):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
When I used the textbook while studying this week, I (select all that apply):

- [ ] Wrote notes on my own paper
- [ ] Wrote notes in the textbook
- [ ] Highlighted the text in the book
- [ ] Printed out online textbook pages
- [ ] Worked out practice problems on my own paper
- [ ] Just read, did not take notes
- [ ] Did problems at the end of the chapter
- [ ] Looked at problems, but did not work them out on paper
- [ ] I did not use the textbook to study
- [ ] Other (please specify): ____________________________________________
- [ ] Other (please specify): ____________________________________________
- [ ] Other (please specify): ____________________________________________
- [ ] Other (please specify): ____________________________________________

Please share comments you may have on your experiences with, opinions on, or suggestions regarding textbooks and their role in a science learning environment.

____________________________________________________________________
The total amount of time I spent studying for (or working on) chemistry the past week was:

- Fewer than 3 hours
- 3-5 hours
- 5-7 hours
- 7-9 hours
- 9-11 hours
- More than 11 hours

When I study for chemistry I do it:

<table>
<thead>
<tr>
<th>Size of study group</th>
<th>Sometimes</th>
<th>Often</th>
<th>Most of the time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alone</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In a group of 2-3 people</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>In a group with more than 3 people</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Only in PLTL groups</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (please specify):</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
On a scale of 1-7, how much does my instructor value the textbook?

- [ ] Does not value the textbook at all
- [ ] 2
- [ ] 3
- [ ] Somewhat values the textbook
- [ ] 5
- [ ] 6
- [ ] Very highly values the textbook

Would you recommend a student coming into general chemistry next semester use a textbook as a study resource?

- [ ] Yes
- [ ] No

Please explain your reason for your answer above.

________________________________________________________________________________________

Thank you for completing the textbook use survey! Your input is very valuable. Good luck with finals, and have a wonderful summer!
APPENDIX I

Textbook use survey, final version

In the last week, I used the following materials when studying for chemistry, and used the resource approximately _______ (select all that apply):

<table>
<thead>
<tr>
<th>I used this material while studying</th>
<th>How often I used this material</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Never</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Resource Type</th>
<th>Option 1</th>
<th>Option 2</th>
<th>Option 3</th>
<th>Option 4</th>
<th>Option 5</th>
<th>Option 6</th>
<th>Option 7</th>
<th>Option 8</th>
<th>Option 9</th>
<th>Option 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Textbook</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student Study Guide</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Powerpoint slides from class</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Notes provided by the instructor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Notes I wrote while in class</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text about the topic from the internet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homework problems from the textbook</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online homework software (ALEKS, Sapling, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Previous exams or quizzes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Something else (please specify)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The following best describes the textbook I use most for studying:

- I use the textbook recommended by the professor.
- I use a different textbook than the one recommended by the professor.
- I do not use a textbook while studying.

If you do not use a textbook for studying, please answer the following question:

I choose not to use a textbook because (select all that apply)

- My instructor does not use the textbook.
- My instructor does a better job explaining concepts than the book.
- Other sources were all that were necessary.
- It was too expensive.
- I found it to not be useful in previous courses.
- The online textbook was difficult to navigate.
- Textbooks are useful in some courses, but not chemistry.

The form of the textbook I use most for studying is

- the physical textbook
- an online version of the textbook
- I do not use a textbook while studying
For the remainder of the survey, answer with respect to the text that you primarily use while studying.

Which of the following best describes my access to the text?

- [ ] I bought a new textbook
- [ ] I bought a used textbook
- [ ] I bought the text along with online homework software
- [ ] I am borrowing the text
- [ ] I am renting the text
- [ ] I use a text from the library
- [ ] I am sharing the text with another student
- [ ] I do not have access to a textbook because I do not feel I need a textbook for the course
- [ ] I do not have access to a textbook because I did not want to purchase a textbook
- [ ] I do not have access to a textbook because I did not want to purchase a textbook due to high cost
- [ ] I have access to a textbook, but choose not to use it while studying
The reason(s) I chose to use a textbook to study is/are (select all that apply):

- To prepare for an upcoming exam
- To prepare for an upcoming quiz
- To review material before seeing it in class
- To review what I have learned in class
- To get a more thorough look at the material than was presented in class
- To review what was not covered in class
- The instructor suggested it
- To review homework problems
- To help answer homework problems
- To practice problems
- I do not use a textbook

Features of the physical or online textbook I used this week (select all that apply):

<table>
<thead>
<tr>
<th>I used this feature</th>
<th>How useful is this feature in helping me understand the concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Not at all useful</td>
</tr>
</tbody>
</table>

---
<table>
<thead>
<tr>
<th>All text/sections in the chapter</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction text</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summary text</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overview text (bulleted list of main points)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text with real-world applications or examples</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Historical references or vignettes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bolded or italicized words</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Section</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>--------------------------------------------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Mathematical equations</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Chemical equations</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Worked out example problems in the chapter</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Non-worked out follow-up problems in the chapter</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>End of chapter problems</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Images or diagrams illustrating the chemical principles</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Graphs or charts</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>
When I used the textbook while studying this week, I (select all that apply):

☐ Wrote notes on my own paper
☐ Wrote notes in the textbook
☐ Highlighted the text in the book
☐ Printed out online textbook pages
☐ Worked out practice problems on my own paper
☐ Just read, did not take notes
☐ Did problems at the end of the chapter
☐ Looked at problems, but did not work them out on paper
☐ I did not use the textbook to study
The total amount of time I spent studying for (or working on) chemistry the past week was:

- [ ] Fewer than 3 hours
- [ ] 3-5 hours
- [ ] 5-7 hours
- [ ] 7-9 hours
- [ ] 9-11 hours
- [ ] More than 11 hours

When I study for chemistry I do it:

<table>
<thead>
<tr>
<th>Size of study group</th>
<th>Sometimes</th>
<th>Often</th>
<th>Most of the time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alone</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>In a group of 2-3 people</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>In a group with more than 3 people</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Only in PLTL groups</td>
<td>[ ]</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
</tbody>
</table>
On a scale of 1-7, how much does my instructor value the textbook?

- Does not value the textbook at all
- 2
- 3
- Somewhat values the textbook
- 5
- 6
- Very highly values the textbook

Would you recommend a student coming into general chemistry next semester use a textbook as a study resource?

- Yes
- No

Because (choose all that apply)
It couldn’t hurt.

It can help explain ideas.

It is useful for practice problems.

I might have done better if I had used it.

I didn’t use it and still did well in the course.

It is a waste of money.

The professor never used it.

The professor’s notes are all that is necessary.

It is not helpful.

Other resources are more helpful.

Tests are based on lecture, not the textbook.