Confirm Responsibility for Knowledge and Value to Retain Undergraduates in Introductory Statistics

By

KUAN XING

B.E., China University of Mining and Technology, China 2006

M.Ed., Anhui Normal University, China 2010

M.Ed., University of Illinois at Chicago, USA 2015

DISSERTATION

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Educational Psychology
in the Graduate College of the
University of Illinois at Chicago, 2020

Chicago, Illinois

Defense Committee:

Theresa A. Thorkildsen, Ph.D., Chair and Advisor

George Karabatsos, Ph.D.

Yue Yin, Ph.D.

Yoon Soo Park, Ph.D., Massachusetts General Hospital and Harvard Medical School

Kirk A. Becker, Ph.D., Pearson VUE

ACKNOWLEDGEMENTS

Many thanks to my dissertation committee members who offered me the endless support and great suggestions for my dissertation research. First, lots of thanks to my advisor, Dr. Terri Thorkildsen (T2) for her great mentorship. Without her mentorship and encouragement, I couldn't make it for my Ph.D. degree. Many thanks to Dr. Yoon Soo Park, another great mentor of mine, who taught me a lot of psychometrics and inspired me in the field of medical education research. Thanks to Dr. Yue Yin, Dr. Kirk Becker, and Dr. George Karabatsos. They are all great researchers and experts. I really appreciate their great feedback. In addition, thanks to our fabulous staff Elise Wilson and Alejandra (Alex) Cantero, who helped me a lot when I had questions about the doctoral program as well as the dissertation process.

I sincerely thank a lot of my colleagues and statistics teachers who I worked with. Thanks to Xue Jiang, Sierra Ryan, Persis Driver, Alex Carstensen, Jill Rothamer-Wallenfeldt, Diane Mitchell, Xiaodan Tang, Qiao Lin, Jung Mi Scoulas, Karen Fong, among others, who supported me and were always happy to chat with me during my doctoral study. I am fortunate to have them as my colleagues and friends. Special thanks to Statistics Senior Lecturer Dr. Dale Embers, who is very supportive and open-minded. Thanks to statistics instructors Mr. Barajas and Mr. Kahle for their great help. Thanks to student volunteers who shared their experience in learning introductory statistics with me.

Lastly, I am in debt with my family and personal friends. My parents Jinying Xing and Luping Chang are always there to support me, physically, financially and psychologically. So are my parents-in-law Qirong Dong and Yuanying Li. Without their support and love, I couldn't finish my doctoral journey. My special thanks to my wife, Feiran Dong, who supported me and sacrificed a lot, and to our daughter, Elin. I am so lucky to have her around. Additionally, many

ACKNOWLEDGEMENTS (continued)

thanks to my U.S. friends, especially Vickie Driver and Bart Shore, who welcomed us and hang out with us a lot in Chicago. They treat us as family members and my wife and I are so fortunate to have their invaluable friendship.

Kuan Xing

TABLE OF CONTENTS

CHAPTER 1 1

Significance of the Study 7

CHAPTER 2 9

Rudimentary Knowledge: Definition and Assessment 11

Achievement Motivation in the Statistics Classroom 32

Commitment to Learning Statistics 45

The Current Study 46

CHAPTER 3 50

Participants 50

Procedures 52

Materials 55

Data Analysis Plan 65

CHAPTER 4 69

Variability in Baseline Commitment to Learning Statistics and Academic Background 70

Variability in Conceptual and Procedural Statistical Knowledge Prime 74

Testing of Differences in Post-Test Commitment to Learning Statistics at the Midterm 78

Testing of Differences in Delayed Post-Test Commitment to Learning Statistics at the Final 81

Testing of Changes on Commitment to Learning Statistics Across the Semester 84

Predicting Persistence Decisions and Perceived Statistical Ability in the Full Responsibility Group 89

Predicting Persistence Decisions and Perceived Statistical Ability in the Knowledge Responsibility Group 93

Estimated Retention Rate 95

CHAPTER 5 97

REFERENCES 102

APPENDICES 117

APPENDIX A 118

APPENDIX B 121

APPENDIX C 132

APPENDIX D 135

APPENDIX E 147

APPENDIX F 151

TABLE OF CONTENTS (continued)

APPENDIX G 154
VITA 155

LIST OF TABLES

TABLE PAGE

TABLE I DEMOGRAPHIC INFORMATION 51

TABLE II FREQUENCIES OF REPORTED MAJORS........ 51

TABLE III LIST OF RESEARCH AND COURSE ACTIVITIES BY RESPONSIBILITY GROUP 53

TABLE IV SAMPLE SIZE AND INTERNAL CONSISTENCY FOR COMMITMENT TO LEARNING STATISTICS 56

TABLE V SAMPLE SIZE AND INTERNAL CONSISTENCY FOR THE CONCEPTUAL AND PROCEDURAL STATISTICAL KNOWLEDGE PRIME 59

TABLE VI SAMPLE SIZE AND INTERNAL CONSISTENCY FOR THE UTILITY VALUE OF STATISTICS QUESTIONNAIRE 63

TABLE VII DESCRIPTIVE STATISTICS OF BASELINE COMMITMENT TO LEARNING STATISTICS 70

TABLE VIII NORMALITY OF BASELINE COMMITMENT TO LEARNING STATISTICS 71

TABLE IX DESCRIPTIVE STATISTICS OF PARTICIPANTS' ACADEMIC BACKGROUND 73

TABLE X PERCENTAGES OF COMPLETED HIGH SCHOOL MATHEMATICS/STATISTICS COURSES 73

TABLE XI CORRELATIONS BETWEEN CONCEPTUAL STATISTICAL KNOWLEDGE AND RESPONSIBILITY TO USE CONCEPTUAL STATISTICAL KNOWLEDGE........ 76

TABLE XII CORRELATIONS BETWEEN PROCEDURAL STATISTICAL KNOWLEDGE AND RESPONSIBILITY TO USE PROCEDURAL STATISTICAL KNOWLEDGE 77

TABLE XIII WITHIN-SUBJECTS CORRELATIONS OF COMMITMENT TO LEARNING STATISTICS ACROSS THE SEMESTER........ 85

TABLE XIV SUMMARY OF RAW SCORES AND T SCORES IN INTRODUCTORY STATISTICS EXAMS 89

TABLE XV DESCRIPTIVE STATISTICS OF PERCEIVED RESPONSIBILITY: THE FULL RESPONSIBILITY GROUP 90

TABLE XVI PREDICTORS OF PERSISTENCE DECISIONS AT THE MIDTERM 91

TABLE XVII DESCRIPTIVE STATISTICS OF PERCEIVED RESPONSIBILITY: THE KNOWLEDGE RESPONSIBILITY GROUP........ 93

LIST OF FIGURES

FIGURE PAGE

Figure 1. Theoretical framework. 2

Figure 2. Dotplot of participants' attitude toward statistics by responsibility groups. 72

Figure 3. Estimated marginal means of persistence decisions and perceived statistical ability. 86

Figure 4. Estimated marginal means of success caused by ability, effort, easy task, and good luck.......... 87

Figure 5. Estimated marginal means of failure caused by low ability, little effort, difficult task, and bad luck.......... 88

SUMMARY

Mastering statistical knowledge and cultivating statistical thinking is crucial for success in a variety of disciplines. Learning statistics is challenging for many undergraduates who do not major in statistics (Zieffler et al., 2008). Their commitment to learning statistics in their statistics courses is an important factor for their future academic attainment. This project was focused on exploring why undergraduates should learn statistics by identifying three types of responsibility–responsibility to use conceptual statistical knowledge, responsibility to use procedural statistical knowledge, and responsibility for perceiving the utility value of statistics. To better understand these 3 components of students' commitment to learning statistics in an introductory statistics course, participants were randomly assigned to either a "full responsibility," "knowledge responsibility," or "undefined responsibility" group. Participants in the full responsibility group completed three responsibility primes – one focused on responsibility for using rudimentary conceptual statistical knowledge, one focused on responsibility for using rudimentary procedural statistical knowledge, and a third focused on responsibility for imagining the perceived utility value of statistics. Participants in the knowledge responsibility group completed only the knowledge primes that focused on responsibility for using rudimentary conceptual statistical knowledge and procedural statistical knowledge. Participants in the undefined responsibility group completed none of those responsibility primes. All participants reported their commitment to learning statistics at the beginning, in the middle, and at the end of the semester.

Results showed at baseline participants from the 3 responsibility groups reported similar commitment to learning statistics. Reportedly, participants in the full responsibility and the knowledge responsibility group were more familiar with procedural statistical knowledge than conceptual statistical knowledge. Participants in the full responsibility group rated their belief about the importance of conceptual/procedural statistical knowledge independently from their scores in those knowledge tests but it was not the case for those in the knowledge responsibility group. At the end of the semester, participants in the full responsibility group tended to report stronger belief about persistence decisions and more likely attributed their statistics success to their effort while their counterparts in the knowledge responsibility group tended to report stronger belief about their statistical ability, both compared with participants in the undefined responsibility group. Across the semester, *all* participants tended to report weaker persistence decisions while they perceived stable statistical ability in learning introductory statistics, which were consistent with previous study (Schau & Emmioglu, 2012). In addition, when reporting their attribution beliefs across the semester, participants tended to report stronger beliefs about success caused by their ability and weaker beliefs about success caused by their effort; they tended to report stronger beliefs about their statistics failure caused by difficult task.

The findings from this research indicated the benefit of exploring undergraduates' responsibility for learning statistics. Participants could distinguish their conceptual statistical knowledge from their procedural statistical knowledge (Star, 2005) as well as relate the value of statistics to their daily lives (Hulleman et al., 2010). Based on my study, almost all participants completed their introductory statistics course even the course content was difficult for many of them. The pattern of their attributional beliefs across the semester could function as adaptive strategies for their retention in their statistics class (Weiner, 1979) by positively holding the belief about their statistics success caused by their ability and effort as well as admitting the difficulty of the course content when they face the challenges.

Confirm Responsibility for Knowledge and Value to Retain Undergraduates in Introductory Statistics

CHAPTER 1

INTRODUCTION

Mastering statistical knowledge and cultivating statistical thinking is crucial for success in a variety of disciplines, including natural sciences, engineering, business, and social sciences. However, statistics courses are challenging for many undergraduates who do not major in statistics (Zieffler et al., 2008). There are differences in the way students view their responsibility for learning statistics which may impact their commitment in statistics courses. Their commitment to and decision on staying in statistics courses are crucial for their future academic attainment. To understand possible reasons for their commitment and retention decisions, previous studies have traced reasons back either to students' preparedness in their basic knowledge or their motivational beliefs about the perceived value of statistics. Focusing on both reasons, I invited undergraduates who were enrolled in introductory statistics courses to report their responsibilities for learning statistics. Their commitment to learning statistics was reported at the beginning, middle, and end of the semester.

Three types of responsibility were investigated: responsibility to use conceptual statistical knowledge, responsibility to use procedural statistical knowledge, and responsibility for imagining the utility value of statistics. Figure 1 represents my proposed theoretical framework. Conceptual and procedural statistical knowledge are two crucial types of knowledge (Anderson, 1982) in statistics. Undergraduates' perceived responsibility to use those two types of knowledge (two ovals in the upper-left corner) may predict the extent to which they will commit to learning statistics. Findings in support of expectancy-value theory haves indicated that the amount of value undergraduates put on statistics may impact their commitment to such learning (Wigfield & Eccles, 2000). Individuals' responsibility for imagining the importance (utility value) of statistics (the oval in the lower-left corner) may predict their commitment to learning statistics. Multiple components such as perceived persistence, attributions of success and failure in statistics, and attitudes toward statistics have been investigated in isolation, but have not been compared well enough to understand the qualities of individuals' commitment to learning statistics (not shown in Figure 1). Investigations into those components can help to better understand the similarities and differences of students' commitment to learning statistics.

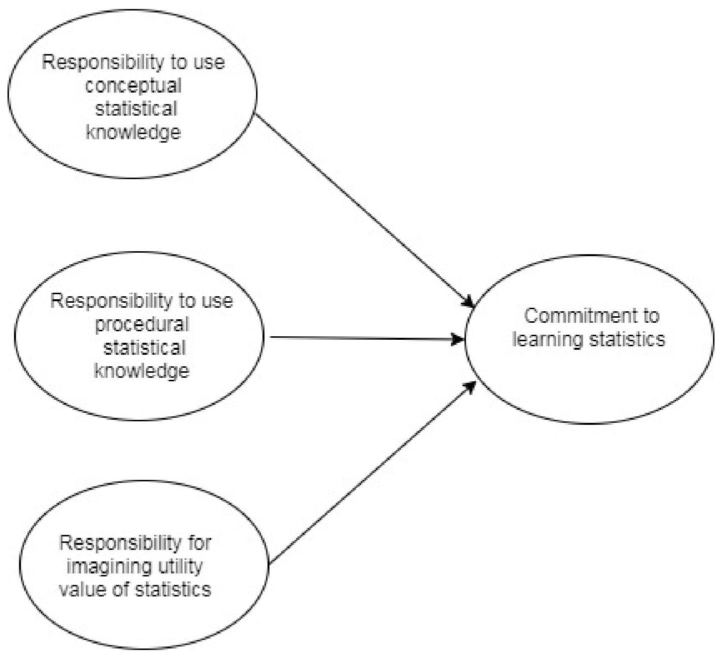

Figure 1. *Theoretical framework*.

Statistics is viewed as an essential but difficult discipline for many undergraduates to learn (Zieffler et al., 2008). Various academic programs require undergraduates to take at least one introductory statistics course in fulfilling their degree requirements. Ideally, students should be aware of their responsibility for learning statistics, which is often not the case. Lack of responsibility may lead to unpreparedness and failure in course performance or minimum effort to study course materials. Therefore, it is crucial for researchers and educators to better understand students' sense of their responsibility for learning statistics. By perceiving a greater responsibility, statistics learners may feel more committed to learning the subject matter, and that will lead to better performance and mastery of statistical literacy.

The statistical needs of scientific research, industry, and governance are rapidly growing in the 21st century (Brown & Kass, 2009). Students across different disciplines in post-secondary education are required to master statistics literacy (delMas, R., Garfield, J., Ooms, A., & Chance, 2007). Their success in introductory statistics courses at the undergraduate level predict their future academic achievement as well as their intentions to pursue advanced degrees in majors such as natural science, psychology, sociology, education, and economics among others (e.g., Freng, Webber, Blatter, Wing, & Scott, 2011).

Statistics education, especially the introductory statistics course in post-secondary institutions nurtures students' scientific thinking and quantitative reasoning as well as scientific investigation skills (Carver et al., 2016). According to Cobb (1992), learners of statistics should be able to understand the data patterns in how the world operates by aggregating data and exploring the variability. Statistics learners (especially those enrolled in non-statistics majors) should be able to develop some statistical investigation knowledge/skills, which would contribute to their own discipline in terms of conducting research or enhancing practice (Horton, 2015). Ultimately, undergraduate students who are equipped with statistical expertise can apply statistical knowledge, understand scientific research design, interpret data analysis results, and make predictions and decisions based on collected data and information (Cobb & Moore, 1997;

Wild & Pfannkuch, 1999). Unfortunately, the study of statistics seems to challenge a lot of statistics learners.

On one hand, statistics learners are required to grasp many statistical concepts and principles which are abstract (Garfield & Ahlgren, 1988; Sotos, Vanhoof, Van den Noortgate, & Onghena, 2007). The nature of such abstraction makes it difficult for many learners to understand basic statistical concepts and principles (delMas, 2004). In addition, researchers have acknowledged that a sizeable portion of statistics learners hold systematic patterns of errors or misconceptions on basic statistics concepts such as probability (Cohen, Smith, Chechile, Burns, & Tsai, 1996). More complex statistical concepts may require a process of repeated learning to fully understand them (Garfield & Ben-Zvi, 2008). Without the formal training and practice in statistics, many students may resist changing any of those misconceptions and fail to grasp the correct statistical conceptual knowledge.

On the other hand, to successfully conduct statistical analysis, statistics learners need to master statistical knowledge about investigative procedures in solving statistics-related problems. The procedural knowledge is mainly focused on problem-solving procedures and related strategies (Anderson, 1982; Wild & Pfannkuch, 1999). In introductory statistics courses, for instance, learners are supposed to be confident in operating statistical symbols, evaluating equations, and solving statistics problems (Schutz, Drogosz, White, & Distefano, 1998), which can be another challenge for them. Since those tasks involve using procedural knowledge, the acquisition of relevant statistical concepts or principles alone may not guarantee the success of conducting a successful solution for the given problem.

One optimal goal for statistics learners to achieve in their introductory statistics courses is to develop their statistical problem-solving ability (Carver et al., 2016). This is especially the case for non-statistics majors who are going to use statistical knowledge in their own disciplines. The underlying cognitive processes used by successful problem solvers were categorized into two basic phases: conceptual representation of the given problem and applying procedures to produce solutions or “products” (Anderson, 1982). Accordingly, successful statistics problem-solvers should be able to use both conceptual and procedural statistical knowledge appropriately and collaboratively. *Conceptual knowledge* enables statistics learners to conceptually represent the problem status in their head and to match it with their knowledge schemata. *Procedural knowledge* enables those learners to use/apply the relevant statistics procedures, to develop possible solutions, and to carry out correct procedures to achieve a successful solution.

When undergraduates enter their introductory statistics courses, many may not have the basic conceptual and procedural statistical knowledge that instructors assume they should have. For instance, some students may have basic conceptual statistical knowledge but are underprepared for procedural statistical knowledge, or vice visa. Learners’ rudimentary knowledge has significantly predicted their future achievement in specific disciplines (Dochy, De Rijdt, & Dyck, 2002; Schwartz, Sears, & Chang, 2007). Both conceptual and procedural statistical knowledge are crucial for understanding the content of introductory statistics courses and succeeding in problem-solving with accurate procedures.

More importantly, undergraduates should be aware of their responsibility for using both types of statistical knowledge: conceptual statistical knowledge and procedural statistical knowledge (de Jong & Ferguson-Hessler, 1996). In Chapter 2, by reviewing the relevant literature I argue that knowledge tests that focus on *both* conceptual statistical knowledge and procedural statistical knowledge offer a better indication of undergraduates’ statistical knowledge than tests focusing on one or the other type of knowledge. I also argue that it is beneficial to ask individuals to reflect on their responsibility to use conceptual statistical knowledge and procedural statistical knowledge.

Still, statistics learners may behave differently in learning statistics because of the potential variability in their motivation to learn statistics. Beliefs about the reasons why they should learn statistics and their expectations for success can play a crucial role. Expectancy-value theory explains that both students' expectation for success in statistics and the value they attach to the course/learning are two important determinants of their motivation to learn statistics (Wigfield & Eccles, 2000). Students' expectancy belief about how well they will perform was one of the strongest predictors of their academic achievement verified immediately or in longitudinal research (Eccles et al., 1983; Wigfield & Eccles, 2000). In the domain of statistics, a meta-analysis study showed a medium effect size regarding the relationship between college students' statistics expectancy and their statistics achievement (Emmioglu & Capa-Aydin, 2012). Achievement-value is one part of this expectancy-value model and is usually defined as a multi-dimensional construct that consists of attainment value, interest value, and utility value (Wigfield & Eccles, 1992). Expectancy and value have been positively correlated in early research (Eccles & Wigfield, 1995). Therefore, I expected to find similar relations in this study.

Lack of achievement motivation may lead individuals to demonstrate a lack of interest in studying statistics and failure to complete their statistics courses. Motivation interventions under the expectancy-value framework have manipulated the value component, and previous studies found a value intervention increased participants' task value and interest and predicted choices of which activities to do (Acee & Weinstein, 2010; Hulleman, Godes, Hendricks, & Harackiewicz, 2010; Wigfield, Tonks, & Klauda, 2016). By emphasizing the relevance of a learning domain such as psychology to individuals' daily lives, findings from the utility value intervention approach have demonstrated a higher sense of responsibility and an increase in students' perceptions of the utility value of the skills they are learning (Hulleman et al., 2010). In Chapter 2, I elaborated on how I applied a utility value prime to investigate statistics learners' perception on the relevance of statistics to their daily lives, and how that perceived responsibility influenced students' commitment to learning statistics.

Significance of the Study

Mastering statistics is crucial for undergraduates across different academic programs to succeed in their programs and their careers. Still, statistics courses are challenging for many non-statistics majors. Retaining those students in introductory statistics courses is a crucial step for encouraging them to meet their academic goals.

Focusing on only students' knowledge acquisition and performance on statistics course exams is not enough to explain the possible reasons for their retention. Instead, it is helpful to investigate individuals' responsibility in their statistics courses. Undergraduates are "emerging adults" who are supposed to take responsibility in many aspects of their lives including learning. The fact that many undergraduates feel they are not ready for independence and taking responsibility calls attention to the need for further investigation.

In the context of statistics, previous studies have focused on knowledge preparation or motivation intervention to help students achieve excellence (Emmioglu & Capa-Aydin, 2012; Schwartz et al., 2007). More detailed studies on how undergraduates take responsibility for using statistics knowledge and for imagining a motivational component in statistics learning can help instructors retain students even when the content becomes difficult. My study will add new knowledge about undergraduates' responsibility to use statistical knowledge and to imagine the value of statistics, and how their perceived responsibility may impact their commitment to learning statistics.

Undergraduates' motivation to learn statistics has sometimes declined across the semester (Schau & Emmioglu, 2012). To better understand whether individuals' commitment changes across the semester, I also track individual commitment to learning statistics at the beginning, in the middle, and at the end of the semester. Information from my study can explore whether undergraduates' commitment declines well enough to inform statistics teachers and educators about why they should talk explicitly to their students about learners' responsibilities.

CHAPTER 2

LITERATURE REVIEW

Undergraduate students who have enrolled in their introductory statistics course face multiple challenges in mastering statistical knowledge and skills, given that subject matter in statistics is perceived to be difficult for many undergraduates (Zieffler et al., 2008). It is crucial to explore the reasons why undergraduates remain in introductory statistics courses rather than dropping out. One aspect focuses on individual students' responsibility to use conceptual and procedural statistical knowledge. Applying the knowledge compilation model (Anderson, 1982), I argue that undergraduates should be able to understand the responsibility to use both conceptual and procedural statistical knowledge, after they have completed one conceptual statistical knowledge prime and one procedural statistical knowledge prime. Another aspect focuses on their responsibility for imagining the utility value of statistics. Adapting the expectancy-value model and the utility value intervention (Hulleman et al., 2010; Wigfield & Eccles, 2000), I argue that undergraduates should be able to understand the responsibility for their perceived utility value of statistics by completing an essay about the relevance of statistics to their daily lives. Finally, I argue that the responsibility to use conceptual statistical knowledge & procedural statistical knowledge and to perceive the utility value of statistics will be associated with individuals' commitment to learning statistics. And by tracking individuals' commitment to learning statistics across the semester, I want to confirm if different components of individuals' commitment change over time.

In general, new knowledge is built on existed knowledge. Mastering basic knowledge in a specific field will influence the acquisition of advanced knowledge. Statistical knowledge has different types or levels of quality. Review of the current literature shows that previous studies have predominantly focused on an approach for promoting conceptual statistical knowledge and understanding (e.g., Garfield, 2002; Schau & Mattern, 1997; Tempelaar, Gijselaers, & van der Loeff, 2006). However, the importance of procedural statistical knowledge is somewhat underestimated (Rittle-Johnson & Koedinger, 2005; Star, 2005, 2007). Applying the knowledge compilation model (Anderson, 1982), I elaborate the essences of both conceptual and procedural statistical knowledge. More importantly, it is a crucial step to investigate whether students can realize that they need to be responsible for using both conceptual statistical knowledge and procedural statistical knowledge. In my study, undergraduates were invited to complete one conceptual statistical knowledge prime and one procedural statistical knowledge prime. In each of those primes, they completed a specific knowledge test designed for conceptual statistical knowledge or procedural statistical knowledge. They then self-rated their responsibility to use those two types of statistical knowledge in their introductory statistics course.

The expectancy-value model emphasizes that individuals' expectancy and value beliefs are two important motivational constructs, which may influence academic behavior, performance, or intention (Wigfield & Eccles, 2000). Previous motivation interventions with an emphasis on perceived utility value revealed a valid and less-threatening way to enhance students' academic motivation (e.g., Hulleman & Barron, 2016; Yeager & Walton, 2011). Results showed individuals in the intervention group reported higher interest than their counterparts in the comparison group (Hulleman & Harackiewicz, 2009; Hulleman et al., 2010). I argued that this approach is appropriate for exploring individuals' imagined relevance of statistics to their daily lives in undergraduates' introductory statistics course.

From the literature review, I concluded that cognition and motivation should function together in terms of raising individuals' responsibility for learning statistics. Asking students to confirm their responsibility to use statistical knowledge and for imagining the utility value of statistics may enhance their commitment to learning statistics. Findings from my research can add new knowledge to the literature about what responsibility tasks can prompt students to focus on their commitment to learning statistics. It is beneficial to remind undergraduates about the value of learning statistics by focusing on responsibility to use conceptual and procedural statistical knowledge and to internalize the importance (utility value) of statistics by relating it to their daily lives.

I also reviewed possible components of undergraduates' commitment to learning statistics. Constructs such as persistence decisions, attributions of success and failure in statistics, and attitude toward statistics were selected and explained. By demonstrating different components of individual's commitment across the semester, I argue the results of changes in their commitment is related to the possible reasons for explaining undergraduate retention in introductory statistics course.

Rudimentary Knowledge: Definition and Assessment

To investigate responsibility to use statistical knowledge, I start with the review of rudimentary knowledge in statistics. Rudimentary knowledge (or prior knowledge, background knowledge) is a term within the framework of cognitive psychology and learning sciences. One theoretical assumption is that our new knowledge acquisition is often based on existing knowledge. Rudimentary knowledge is a powerful definition that many researchers in different fields have adopted (Dochy, De Rijdt, & Dyck, 2002). Broadly, an individual's rudimentary knowledge can be defined as "the whole of a person's knowledge" (Dochy et al., 2002, p.267). Rudimentary knowledge reflects an individual learner's existing knowledge structure.

Nonetheless, it is a general term and can be broken down into a set of sub concepts. I highlighted several dimensions that might help in understanding rudimentary knowledge in the literature.

One dimension of rudimentary knowledge distinguishes semantic knowledge from episodic knowledge (Bransford, 1979). Episodic knowledge refers to an individual's experience that is situated in a spatio-temporal manner. This type of knowledge often covers the information and memories that one experiences in the informal, daily life. Semantic knowledge, by contrast, frequently refers to the information and memories in the formal educational settings.

Another dimension of rudimentary knowledge focuses on domain-specific and domain-transcending knowledge. Domain-specific knowledge is one’s knowledge of a specific domain or subject. For instance, one’s knowledge system may consist of mathematics, history, and other domains. In addition, an individual may have a rich knowledge system in mathematics but a poor knowledge system in the history domain. Domain-transcending knowledge is one’s knowledge that can apply across different subjects or domains. An individual’s domain-transcending knowledge partially explains why learners could solve novel problems which are not familiar to them.

Both domain-specific and domain-transcending knowledge are crucial in problem-solving and task performance across different disciplines (Chi, Glaser, & Rees, 1982; Chase & Simon, 1973; Voss, Greene, Post, & Penner, 1983). Domain-specific knowledge is crucial to distinguish between domain experts and novices (Chi, Feltovich, & Glaser, 1981). A domain expert’s knowledge organization to represent the domain-specific problems is much more complex and better organized than a novice learner’s knowledge organization. When problem-solvers encounter novel problems, domain experts can use their prior domain-specific knowledge (expertise) more precisely and effectively to generate a mental representation of the problem and search strategically for the possible solutions compared to their novice counterparts.

Research on domain-transcending knowledge has focused on strategic knowledge in problem-solving settings and revealed that strategic knowledge is necessary for seeking solutions. The "weak method" strategy, or the so-called heuristic strategy is characterized as domain-transcending (Anderson, 1985). Particularly, problem-solvers tend to refer to the general methods when they face novel questions that are not domain-specific (Glaser, 1983). This type of strategic knowledge can be applied in the settings where it does not require much domain-specific knowledge.

Still another dimension of rudimentary knowledge has focused on the *types* of knowledge that can function differently in specific types of tasks (Glaser, 1983). Researchers have emphasized two types of knowledge, conceptual knowledge and procedural knowledge (Anderson, 1982; Hiebert & Lefevre, 1986; Star, 2005). Many research discussions have occurred in the field of mathematics. *Conceptual knowledge* is defined as an individual's knowledge about facts, principles, or rules (Hiebert & Lefevre, 1986). Sometimes, it can be briefly referred to the knowledge about "knowing that" (Byrnes & Wasik, 1991). In addition, some other researchers have proposed that conceptual knowledge should also represent the interrelations of pieces of information or propositional networks in a specific domain (Rittle-Johnson & Alibali, 1999). *Procedural knowledge* is defined as an individual's knowledge about sequenced actions (Rittle-Johnson & Alibali, 1999). Sometimes, it can be briefly referred to the knowledge about "knowing how" (Byrnes & Wasik, 1991). Broadly, individual's procedural knowledge may include knowledge of symbols, algorithms, procedures, step-like skills, and/or productions.

Different types of knowledge may apply for various tasks in a specific knowledge domain (De Corte & Vershaffel, 1987). In statistics, the type of task on conceptual knowledge may require the recognition and application of information that learners have acquired. In contrast, the type of task on procedural knowledge may require the production of information about learners' internalized actions. In addition, these two types of knowledge may be complementary to assist statistics students with their problem-solving process (Rittle-Johnson et al., 2015).

Some undergraduates may acquire procedural statistical knowledge but less likely master conceptual statistical knowledge, or vice versa. In the next section, the knowledge compilation model indicates that both types of statistical knowledge are crucial, and therefore it is important to investigate undergraduates' perceptions about their responsibility to use both types of knowledge.

Knowledge Compilation Model

As a cognitive model for illustrating a cognitive skill acquisition, a knowledge compilation model initially explained how problem-solvers' knowledge functioned during the problem-solving process (Anderson, 1982). The process of problem-solving usually involves two main phases: a conceptual knowledge phase and a procedural knowledge phase. The first phase involves instantiation of conceptual knowledge (sometimes referred to as declarative knowledge, although researchers may argue they are somewhat different). In this phase, problem-solvers may encode the information based on the presentation of the given problem. Problem-solvers may also connect this information conceptually to their pre-stored ideas—their existing knowledge structures or schemata. Here, conceptual knowledge is often represented as propositional networks.

The second phase involves the compilation of procedural knowledge that can direct problem-solvers to perform their knowledge or skills in problem-solving. In Anderson's model, conceptual knowledge about a problem or task can be converted to procedural knowledge. Procedural knowledge is often represented as productions (Anderson, 1982). Each production has a condition that specifies certain circumstance where the production can apply and certain actions (or a set of steps) respond to what to do to solve a problem.

An illustrated example of the problem-solving processes using the conceptual and procedural knowledge of statistics is described briefly here (see detail in Appendix A). Suppose a statistics question asks learners to investigate whether there is a difference between two groups of students (male versus female) on one mathematics standardized test score. Anderson's model can be applied to explain the possible processes that problem-solvers may take. In the first phase, problem-solvers may encode the information from the given question. Problem-solvers may encode key information such as difference, two groups, and mathematics test scores. A search will be conducted to find out the most appropriate concepts and/or associated principles that match the encoded key information. Problem-solvers with sufficient statistics conceptual knowledge may come up with the concepts of inferential statistics, *t*-tests for mean difference comparison, and the independent two group *t*-test. This type of conceptual knowledge helps problem-solvers to generate a mental representation of the problem situation and to search for an optimal solution.

In the second phase, problem-solvers may need to apply their procedural knowledge for producing the appropriate t-test analysis. Suppose that problem-solvers are required to hand calculate their analysis results. Once an optimal (or a close to optimal) solution is developed, problem-solvers with sufficient procedural knowledge may compile their procedural knowledge about the steps of conducting a null hypothesis t-test procedure. It may look like this:

- Step 1: List the H0 and H1 hypotheses.
- Step 2: Confirm an independent, two group *t*-test will be used.
- Step 3: Use the formula to calculate the *t* statistics.
- Step 4: Compare the *t* statistics to the *t*-test significance table (suppose that the significance level is $\alpha = .05$).
- Step 5: Make a decision that there is no significant difference based on Step 4's statistics comparison result.
- Step 6: Finalize the conclusion (interpret) that there is no significant difference between male and female students in their mathematics test scores, $p > .05$. In other words, the score difference between the male and female group is probably due to sampling error.

Anderson's knowledge compilation model highlights some functions and features of conceptual knowledge and procedural knowledge in problem-solving (Glaser, 1991). Different types of knowledge components may help statistics learners to present the given problem in their working memory, search and match their existed knowledge, and compile necessary steps or procedures for succeeding in solving the given problem.

Although Anderson's knowledge compilation model fits the investigation scenario of my study, two limitations should be brought to readers' attention. One potential limitation is that initially the knowledge complication model was applied to a cognitive skill acquisition in well-defined disciplines. Here, I refer to it as a general approach that explains the necessaries to prepare statistics students for developing their essential knowledge (and skills) in statistics problem-solving. Accordingly, this model is supposed to help to justify the cognitive processes of applying procedural knowledge and conceptual knowledge. However, I did not pay attention to specific details on their cognitive process (e.g., record their cognitive protocols) since this was beyond the scope of my study. Examining students' perceptions to distinguish conceptual knowledge from procedural knowledge to reflect on the use of both types of knowledge is the focus in my study.

The second potential limitation is that Anderson's model is constructed for the design of a computer-tutoring system or a programing language (i.e., Adaptive Control of Thought – Rational [ACT-R]). Therefore, his model specifies that conceptual knowledge is limited to facts only, and procedural knowledge is limited to only a (literally) representative form of conceptual knowledge. I argue that this may oversimplify the definitions of conceptual knowledge and procedural knowledge (see the discussion in Alexander, Schallert, & Hare, 1991). Research from mathematics education revealed conceptual knowledge and procedural knowledge are reciprocal (Rittle-Johnson et al., 2015). Therefore, the tasks that I designed or adopted from other studies involved some overlaps between pieces of conceptual knowledge and procedural knowledge for solving problems in the field of statistics. Again, the focus here should be on justifying the perception that both conceptual knowledge and procedural knowledge are important for individual's success in learning statistics.

In sum, Anderson's model explains the potential cognitive processes of acquiring a specific set of knowledge or skills. It is useful to endorse that both conceptual knowledge and procedural knowledge are indispensable and they seem to work collaboratively in different phases regarding the success in problem-solving. In the next two sections, I further discuss the importance of two types of statistical knowledge and how to evaluate those types in the context of an introductory statistics course. The decision-making of selecting the appropriate measurement tools was also justified. Notice the term "rudimentary" was operationalized in my study for two general purposes: First, it was about students' statistical knowledge at the beginning of the semester, hence the use of the term "rudimentary". Second, the knowledge tasks in my research materials should require a set of knowledge that undergraduates should be exposed to in high school or in other areas of their college education.

Conceptual Statistical Knowledge

The importance of conceptual knowledge in human learning has been well documented in the literature (Anderson, 1982; Chi et al., 1981; Novak, 1990). Below are several main findings summarized from the literature. First, the line of research in human problem-solving (and broadly in cognitive science) has demonstrated that encoding the information from the given problem or situation involves problem-solvers' conceptual representation of the problem and the goal to achieve (Anderson, 1982). Second, research on differences of domain experts and novices' knowledge in the cognitive sciences has shown that experts gained more conceptual knowledge and used it more effectively (Chi et al., 1981). Finally, research on scientific concept learning has shown that conceptual knowledge is an essential component of learners' mental representation of their knowledge system (Novak, 1990).

Conceptual knowledge has at least several different definitions (Byrnes & Wasik, 1991; Hiebert & Lefevre, 1986; Rittle-Johnson & Alibali, 1999). In the field of statistics, rudimentary conceptual knowledge can be defined as the knowledge to-date about statistical concepts and principles, interrelations of statistical concepts, and/or hierarchical propositional networks. For example, knowledge about the concept of "distribution" is one piece of statistics conceptual knowledge. Within that, an information of the connection between "normal distribution" and "central tendency"—for example, the statement of "a normal distribution has its measures of central tendency (mean, median, and mode) that are all equal" is another piece of conceptual knowledge (interrelations of two concepts). Finally, when distribution, normal distribution, and central tendency were put together, a hierarchical organized map (e.g., Novak & Gowin, 1984) can map out a knowledge structure that consists of different concepts and their interrelated propositional networks. Such a hierarchical map can represent a complex level of conceptual knowledge.

It is well known that many researchers in statistics education have called for the promotion of students' conceptual knowledge in the last few decades. There are several reasons for this call for education reform. First, it has been problematic in the past that many statistics learners predominantly focused on mechanical knowledge in statistics (Garfield, 1995). Students who lack conceptual understanding of statistical ideas may still get some test items correct. However, since they misunderstand the statistical concepts, they may fail if there is a minor change in the statistical task. In other words, the acquisition of statistics knowledge among those students happened at the surface level (de Jong & Ferguson-Hessler, 1996). Second, a sizeable number of statistics teachers may hold a simplified view that it is sufficient to teach their students the techniques as long as those students can correctly respond to statistics test questions (Smith, 1987). Finally, some inappropriate instructional designs in statistics classes may hinder students' learning of statistics conceptual knowledge (Moore, 1997).

Selecting measures of conceptual statistical knowledge. Conceptual statistical knowledge can be measured in different formats. Broadly, the formats of the measures can be divided into traditional tests and a concept mapping technique. A line of research in statistics education has focused on the statistical concepts and "statistical literacy" (delMas et al., 2007; Tempelaar et al., 2006). Those researchers applied multiple-choice items (traditional format) that asked about students' conceptual knowledge and understanding in their introductory statistics courses. Another line of research has focused on adapting concept mapping techniques (Schau & Mattern, 1997). The application of the concept map is seen as relatively new format for knowledge assessment (Haapala, Pietarinen, Rautopuro, Valtonen, & Vaisanen, 2002; Novak, 1990; Roberts, 1999).

The traditional test of conceptual knowledge. Two examples of traditional assessment of statistics conceptual knowledge were introduced. A group of statistics education researchers developed a tool called Comprehensive Assessment of Outcomes in Statistics (CAOS) for assessing students' conceptual knowledge at the end of their introductory statistics courses. Forty multiple choice items were given to a national sample of undergraduates from four-year university/college as well as two-year technical/community college on 10 domains including data collection design, descriptive statistics, and probability among others (delMas et al., 2007). Pretest and post-test versions of the assessments were conducted and participants reported that they gained or remained steady in their knowledge level in some domains and had more difficulty in the rest of the assessment domains. Students did not demonstrate a good understanding about probability, sampling variability, and inferential statistics.

Statistical reasoning assessment (SRA) investigated two statistics knowledge domains: correct reasoning and misconceptions about reasoning (Garfield, 2002, 2003; Tempelaar et al., 2006). The SRA consists of 20 multiple-choice items and measures 16 categories in statistics (eight for correct reasoning and another eight for incorrect reasoning). However, validity studies have shown the usage of the SRA in statistics class had no significant relation to students' statistics course performance (Tempelaar et al., 2006).

Studies that used either CAOS or SRA did not show a significant effect of the pretest and post-test assessment outcome on statistics performance. Developers of the CAOS assessment (delMas et al., 2007) stated that the test was made intentionally to investigate students' conceptual understanding instead of asking about statistical procedures. Next, we will turn to an alternative assessment format in measuring conceptual knowledge—the concept map.

The concept map test of conceptual knowledge. Many researchers have developed specific knowledge measurement tools for assessing the type of conceptual knowledge (Novak, 1990). Among those measurement tools, concept mapping is a widely used technique to evaluate the meaningful relationships between concepts in the form of propositions (Novak & Gowin, 1984).

The use of the concept map has shown that a higher score of concept map task predicted better student achievement among English learners (Chularut & DeBacker, 2004), science learners (Novak, 1990), and statistics learners (Haapala et al., 2002; Roberts, 1999; Schau & Mattern, 1997), among others. A meta-analysis review on the effect size of conceptual knowledge on achievement revealed an average effect size of 0.34 for the relation between the application of conceptual knowledge intervention and students' later performance (Nesbit & Adesope, 2006). Nesbit and Adesope reviewed many studies from different subjects (including statistics). One subfinding shows the average effect size is around 0.77 at the undergraduate level courses. Another subfinding demonstrates the effect size rises from 0.34 to 0.52 when students use concept maps as assessment tools instead of learning tools in general science subjects.

Concept mapping is supposed to tap into learners' cognitive structure and make their knowledge organization about specific topics explicit for learning and assessment (Novak & Gowin, 1984). It can serve different purposes. When using it as a testing tool for scoring students, researchers recommended following several scoring criteria: "(1) hierarchical levels of conceptual knowledge; (2) the extend of conceptual differentiations; and (3) the extend of integrative reconciliations" (Novak & Gowin, 1984, p. 107).

The first scoring criterion is that conceptual knowledge can be evaluated based on levels of hierarchy. Knowledge is more meaningful when more layers of the relationship can be expressed. Scoring will focus on the number of valid levels and should be given multiple points for each level. Second, the extend of conceptual differentiations refers to the dimension of propositions. Concepts are assumed to be connected with propositional links. Scoring will focus on the number of relationships that are valid, and a total score can be generated from counting those valid interrelations. The third dimension is the cross-links, representing valid relationships between two distinct segments of the concept hierarchy. They refer to the extent of integrative reconciliations. According to Novak and Gowin (1984), the number of correct cross-links could be doubled by the number of hierarchical levels. Lastly, sometimes researchers use the number of specific examples as another dimension if applicable.

In the field of statistics, existing studies applied the scoring rules for concept maps that were very similar comparing to the Novak and Gowin's (Haapala et al., 2002; Roberts, 1999; Schau & Mattern, 1997). For instance, college students in their statistics class were asked to create their concept maps for a specific topic—statistical inference or for a general topic—introductory statistics (Haapala et al., 2002). Multiple dimensions of their assigned concept maps were evaluated including terms, links, propositions, hierarchies, and examples based on a 0 – 6 scale (example was based on a 0 – 2 scale). The total score was then calculated by summing up all dimensional scores.

One practical problem related to the concept mapping task is that there is still a debate on whether to choose creating concept maps or to fill in pre-created concept maps (Schau & Mattern, 1997; Ruiz-Primo & Shavelson, 1996; Yin & Shavelson, 2008). On one hand, a fill-in concept map is more reliable and takes less time for students to complete; but it may be at risk from low validity measures (Yin & Shavelson, 2008). On the other hand, the concept map for creating links may have more variance in its outcome and requires more time and detailed instructions but it may tap into a higher internal validity in terms of score explanation. In practice, when a concept map with fill-in questions is well constructed, it is reasonable to apply a fill-in testing form of concept map to evaluate students' conceptual statistical knowledge (Schau & Mattern, 1997).

Another problem associated with the use of concept mapping is that relevant research has shown inconsistent results for the effect of this tool to evaluate the pre- and post-test differences (i.e., knowledge performance differences). Systematic literature review showed that there was a nonexistent-to-moderate effect of concept map testing scores on students' achievement (Ruiz-Primo & Shavelson, 1996). Further analysis found that the effect of concept map test may be more effective if the following achievement test taps into students' higher-level thinking (Wallace & Mintzes, 1990). In statistics, previous research indicated that a couple of dimensions from concept mapping on statistical conceptual knowledge predicted statistics learners' later achievement in statistics problem-solving (Roberts, 1999). However, in my study I did not make any performance prediction using the concept mapping task since the selected task had nothing to do with their curriculum.

In sum, I selected the concept mapping technique to construct my knowledge test regarding the conceptual statistical knowledge for several reasons. First, it could help students to explore and self-evaluate their rudimentary conceptual knowledge in statistics (Roberts, 1999). Second, previous studies indicated that by doing this, students may become aware of the importance of their rudimentary knowledge for impacting their learning on new statistical knowledge. With the received feedback, statistics learners may modify the misconceptions they already have and deepen their understanding of correct conceptual knowledge. Finally, evaluation studies on the fill-in concept map test have generally shown good reliability indices (Ruiz-Primo, Schultz, Li, & Shavelson, 2001). My test, which adopted the fill-in concept map task, may be easier for students to answer and require a smaller amount of cognitive load than those with the constructing format. This is especially suitable for many undergraduate students who have just started in their introductory statistics courses. The key concept of "probability" was chosen since it was a central topic in elementary statistics (Cobb, 1992; delMas et al., 2007).

Procedural Statistical Knowledge

As discussed above, conceptual knowledge can be defined as the knowledge about knowing/understanding the propositional networks including concepts and their interrelations. However, acquisition of conceptual knowledge does not guarantee success in applying such knowledge in problem-solving. In addition, it is not always the case that this is how students start to learn basic knowledge in a specific domain. Smith (1987) argued that understanding the statistical concepts and being able to carry out correct statistics procedures are separate knowledge domains. He illustrated this by an example showing that conceptually knowing what the definition of a *t*-test is does not predict solving a *t*-test question correctly, or vice versa.

According to Anderson (1982), procedural knowledge (phase) in problem solving involves compilation of knowledge embodied in the procedures for performing tasks/solving problems. Similarly, solving problems in statistics context requires not only conceptual knowledge but also procedural knowledge. Arguably, statistics is a mathematical science (Moore & Cobb, 2000). Therefore, in many situations it borrows a mathematical language. Procedural statistical knowledge (albeit rudimentary), such as arithmetic calculation and algebra expression, has shown the significant effect on predicting statistics learners' achievement and success in their statistics courses (Galli, Chiesi, & Primi, 2011; Neimark & Burdman, 1980; Schutz et al., 1998). Even though a lot of introductory statistics courses do not require heavy mathematics skills, understanding and mastering procedural knowledge may help students see the essence from the surface. For instance, learners who are equipped with some solid procedural knowledge about mathematics symbols/formula, calculation and expression could achieve a so-called "procedural flexibility," which predicted an increase on high-order knowledge and skills and on "generating problem solutions similar to content experts"(Maciejewski & Star, 2016; Star, 2005).

When searching the literature, there are fewer studies on procedural statistical knowledge in the field of statistics education compared to their counterparts on conceptual statistical knowledge. One possible reason is that in the past several decades, criticism on mechanical learning and procedural instruction in the field of teaching statistics has been widespread (Chance, 1997; Cobb, 1992). However, previous studies have not explicitly distinguished the type of knowledge with the quality of knowledge (de Jong & Ferguson-Hessler, 1996). Previous critics on procedural knowledge (e.g., Hiebert & Lefevre, 1986) asserted that procedural knowledge is a type of *isolated* knowledge. In fact, it refers to a phenomenon of poor quality of procedural knowledge, or rote procedural knowledge—the knowledge that is mechanically memorized by learners. The simple procedural calculation (without conceptual understanding) is an example of that surface level of the procedural knowledge acquisition. A recent review on defining procedural knowledge has indicated that procedural knowledge can be meaningful and flexible (Baroody, Feil, & Johnson, 2007; Star, 2005, 2007).

Another possible reason is that experienced learners or domain experts have usually automated their procedural knowledge, which makes it more difficult to conduct relevant research (Chi et al., 1982). For example, domain experts may be good at incorporating a good deal of procedural knowledge while novices may lack procedural knowledge. Worked examples can be an effective tool (scaffolding) for helping new learners to master procedural knowledge (Rittle-Johnson & Koedinger, 2005).

In the context of statistics problem-solving, procedural statistical knowledge as one crucial type of knowledge is an integral part of the subject matter in the field. Arguably, underestimating the importance of procedural knowledge in statistics may have a negative influence on students' competency in applying statistical knowledge because they may lack the ability to recall or "produce" the step-by-step procedural knowledge (Anderson, 1982). In an even worse case, statistics learners may rely on their mechanical memory on incomplete or incorrect procedures. Therefore, further understanding of procedural statistical knowledge is beneficial for statistics learners. Lack of procedural knowledge may lead undergraduates to a failure on statistics calculations or procedures which are crucial for solving statistics-related problems.

In my study, I choose to focus on two categories of procedural statistical knowledge. The first category is described as the knowledge of statistics/mathematics symbols and rule of calculations, i.e., arithmetic and algebraic skills (Hiebert & Lefevre, 1986). Students are supposed to be able to recognize and use different symbols and calculation rules properly (Byers & Erlwanger, 1984). The second category is the knowledge about carrying out real-world statistical analysis problems, e.g., calculation of probability in the word problems of statistics.

In the context of statistics, in order to read a statistics textbook or solve a statistics problem which usually involves use of symbols and calculations, statistics students need to master statistics/mathematics symbols and calculation rules. For instance, one statistics task requires students to calculate the weighted means (e.g., GPAs). On a 0 – 4 grade scale, one's final grades for course X, Y, and Z are 4.0, 3.0, and 2.0. The course credits for course X, Y, and Z are 3, 4, and 2. In order to complete the task, students need to know the procedural knowledge as follows: (1) calculating the total credits for each course, 3 + 4 + 2 = 9; (2) calculating the total credits, 4 * 3 + 3 * 4 + 2 * 2 = 28; and (3) calculating the GPA using the total credits divided by total credits, 28 / 9 = 3.11 (Pollatsek, Lima, & Well, 1981). Similarly, if the question is formatted using symbols, e.g., GPA = $\frac{\sum_{i=1}^{n} x_i a_i}{n}$ = ?, students also should know this it is the same question which is expressed by the statistical language/symbols.

The importance of mastering procedural knowledge in statistics is at least two-fold. First, knowledge on the use of statistics symbols and rules will increase statistics learners' skill automation and reduce learners' cognitive load (Schneider & Stern, 2010). Anderson (1982) discussed the automation of cognitive knowledge/skill: An experienced problem solver could unconsciously solve a problem by intriguing their automated procedural knowledge with minimal effort. In the context of statistics, students who gained experience in using procedural statistical knowledge for worked examples may enhance their ability in problem-solving and demonstrate more effective knowledge transfer (Paas, 1992).

Second, acquisition of procedural knowledge based on its conditional (or goal-specific) application settings will increase problem-solvers' ability to organize knowledge in a meaningful way. For instance, procedural flexibility is an indicator of problem-solvers' ability in problem-solving (Star, 2005). In the beginning, learners may acquire subprocedures which help them solve simple tasks/problems (Hiebert & Lefevre, 1986). Gradually (with practice), subprocedures may be embedded in other more complicated, macro-level procedures – superprocedures among those experienced learners. Problem-solvers who have mastered the knowledge of subprocedures and superprocedures are more likely to solve domain-specific problems in novel problem settings.

In the scope of my study, the focus on the procedural statistical knowledge would be appropriate usage of symbols/operating rules as well as knowledge about carrying out statistics procedures (e.g., probability calculation). Notice that (1) the content of the test is not necessarily representative of the scope of students' knowledge, and (2) it is on purpose that the content of the procedural statistical knowledge has overlapped (e.g., probability) with the conceptual statistical knowledge. This reflects the argument which I tried to make previously: conceptual knowledge and procedural knowledge should not be in a hierarchical order. Instead, successful learners in statistics should be able to appreciate both types of statistical knowledge.

Selecting measures of procedural statistical knowledge. As mentioned before, two categories of procedural statistical knowledge were selected as the content of the procedural statistical knowledge test. The first category was focused on arithmetic and algebraic knowledge, especially use of symbols, linear expression, and calculation rules. This is the prerequisite knowledge component for introductory statistics (Galli et al., 2001). The second category was focused on the knowledge of carrying out statistics procedures. The required procedural knowledge is to be able to recognize the statistics formula and correctly carry out the procedures of data input, calculation, and producing the final result, e.g., for two independent events, the probability that two events happen at the same time is P (AB) = P (A) * P (B) = (final answer).

Tests of arithmetic and algebraic knowledge. One dimension of measuring procedural knowledge in statistics is the knowledge of the usage of statistical/mathematical symbols and calculation rules. The symbols can be viewed as a form of statistical (or mathematical) language (Byers & Erlwanger, 1984). They represent a type of symbolic understanding and are useful in quantitative computing and reasoning (Skemp, 1982). The calculation rules included how to compile numbers and symbols, solve linear equation, and produce appropriate statistics (e.g., mean). Procedural knowledge of arithmetic and algebra serve as rudimentary knowledge for solving statistical problems (Galli et al., 2001; Neimark & Burdman, 1980).

Whether in statistics textbooks or in statistics research, the procedural knowledge about recognizing statistical symbols and carrying out arithmetic/algebraic computations is explicitly required (in statistics textbooks, see Spatz, 2011; in statistics studies, see Galli et al., 2001; Schutz et al., 1998). The testing format on procedural knowledge about statistical symbols and calculation rules usually includes multiple choice items and open-ended items (e.g., hand calculations). The specific tasks in open-ended test questions usually require students to generate a numerical answer or to show the steps by appropriately recognizing symbols and applying the formula, algorithm, or computation. The multiple-choice items do not allow for evaluating the procedures which students may use to solve the problems, which is a drawback of such a test format.

The scoring criterion for multiple choice items on procedural knowledge is binary, either correct or incorrect (Galli et al., 2001; Smith, 1987). Sometimes researchers may count the number of errors that students make (Allwood & Montgomery, 1981). The scoring criterion for open-ended questions can focus on the correctness of specific steps for solving the problems. Their responses can be scored using a partial credit scoring method. For instance on a 0 – 2 scale, score 2 represents fully correct, 1 represents partially correct, and 0 represents incorrect. In my study, I applied the partial credit method for scoring the open-ended questions on procedural statistical knowledge.

Tests of procedural knowledge in solving probability problems. Another category of procedural knowledge is focused on how well students can carry out statistics procedures for solving statistics problems. Correct application of procedural knowledge in problem-solving can generate accurate solution/production (Anderson, 1982). The format of testing usually involves word problems which require students to carry out appropriate statistics procedures (calculations) based on the extracted information and to get the correct result. The literature has recorded the error of application (knowing the rule but not able to correctly apply that rule) among statistics learners (Kahneman, Slovic, & Tversky, 1982).

The criteria for scoring word problems in statistics include two rules: The first rule focuses on the binary score: correct/incorrect. The second rule concerns the accuracy in carrying out statistics procedures. Some students may experience difficulty when transferring information from word problem to statistics language (Cobb, 1992). It is also not uncommon for some learners to get the final answer correct but demonstrate incorrect procedures in solving statistics problems. The format of the test that requires student to demonstrate not only the final answers but also the step-by-step procedures on solving those problems is crucial for informing student learning and promoting their problem-solving ability (Gelman, 2005). I combined those two rules in my study, which can reflect the essentials of procedural statistical knowledge.

Responsibility to Use Conceptual and Procedural Statistical Knowledge

In the United States, the entry-level statistics course serves several key functions for developing statistics students' (1) appropriate understanding of statistical principles and concepts, (2) appropriate understanding of statistical procedures and algorithms, (3) appreciation of statistical knowledge and thinking, and (4) intention to take advanced statistics courses (Allaga et al., 2005). The first two learning goals have emphasized that conceptual statistical knowledge and procedural statistical knowledge are crucial for students to succeed in their introductory statistics class.

More importantly, students' perceptions of those types of statistical knowledge may be problematic. Some students may hold an ill-structured view that statistical knowledge is all about remembering a set of procedures. Accordingly, those who merely memorize statistical techniques may fail to solve real-world statistical problems. Therefore, statistics learners must acquire conceptual statistical knowledge. In fact, teaching statistics in the introductory statistics course has changed in the last several decades from a focus on teaching specific techniques to a focus on teaching for understanding important concepts and reasoning with those concepts (Cobb, 1992; Garfield, 1995; Smith, 1987).

Other students may be aware of conceptual statistical knowledge but they underestimate the importance of procedural statistical knowledge. Conceptual knowledge itself may be not sufficient for statistics students to successfully solve statistics problems. In the practice of teaching, some statistics instructors have made efforts to design their courses and learning materials for enhancing students' understanding of conceptual statistical knowledge. However, students who were taught to predominantly focus on conceptual statistical knowledge may not perform well on solving statistics problems. Undergraduates who were taught conceptually in statistics made more errors in their statistics performance when compared to their counterparts who were not taught in that way (Smith, 1987).

To put them together, it is important to investigate whether undergraduates may feel responsible for the acquisition of conceptual statistical knowledge and procedural statistical knowledge. It is acknowledged from the literature that these two types of statistical knowledge may serve different cognitive functions (Anderson, 1982; Byrnes & Wasik, 1991), and they may interrelate with each other (Rittle-Johnson et al., 2015). Conceptual knowledge reflects how learners order and organize their experiences in learning. Procedural knowledge reflects the means to an achievement goal that learners can use and apply. However, it does not mean learners are aware of both types of statistical knowledge.

Asking undergraduates about their responsibility to use statistical knowledge alone may be somehow subjective. For the design of rudimentary statistical knowledge primes, a concept map task was administrated for conceptual statistical knowledge and a set of word problems was administrated for procedural statistical knowledge. After those tests, detailed answer and feedback were provided. More importantly, students were also reminded to report their perceived responsibility to use such statistical knowledge and to connect rudimentary statistical knowledge with new knowledge which will be taught in their introductory statistics courses. A Know-Want (to know)-Learned framework (K-W-L framework) will be adopted for enhancing the priming effect (Ogle, 1986; see detail in Chapter 3). The questionnaires on responsibility to use conceptual and procedural statistical knowledge are introduced in Chapter 3.

Achievement Motivation in the Statistics Classroom

Undergraduate students in their statistics class are oriented by their goals or motives for learning statistics. Achievement motivation, which employs such goals/motives, plays an important role in students' learning and their academic performance. Individual variations in achievement motivation may affect learners' learning activities such as their choice of future learning tasks and their learning performance (e.g., academic attainments such as course grades). A long-standing achievement motivation theory, expectancy-value theory is discussed in this section for exploring potential reasons on individual differences in the context of statistics classroom (Wigfield & Eccles, 2000).

Expectancy-Value Model

Expectancy-value theory explains why students may differ in their academic behavior and their academic performance (Wigfield & Eccles, 2000). Individuals' expectations of how well they can perform an upcoming task and the value they attach to the task are two important determinants of their behavioral choice and the level of their achievement. In the education field, students' expectations for success on tasks address their self-estimations of their abilities to perform the incoming tasks and assignments. The value of the task is determined jointly by the characteristics of the task and students' personal needs or values (Eccles et al., 1983).

Studies have shown that students' expectancy for success in their schools or classrooms influenced their academic behaviors and future choices (Feather, 1988; Wigfield & Eccles, 2000). The aforementioned term "expectancy" was defined as students' expectations on how well they will do in an upcoming task or assignment in one specific setting. The level of their expectancy may be influenced by individual anticipated ability on the task and the expected difficulty of the task. The emphasis on the term "expectancy" in the expectancy-value model is that the individual's beliefs about their success is future-oriented. In theory, it is differentiated from a line of research which conceptualizes this construct as self-efficacy, which is present-oriented (Bandura, 1997). The emphasis on the term "self-efficacy" is that the individual's belief is evaluated based on what she/he can do. However, in practice, existing findings from empirical studies indicated these two constructs are highly correlated and the expectancy-value model tends to include both types of expectancy (e.g., Eccles & Wigfield, 1995; Hulleman et al., 2010).

In the context of statistics, students' expectancy is conceptualized as their expectation on how well they will do in their statistics courses. The expectation is based on their beliefs about the excellence they will achieve. Their self-ability beliefs may relate to their prior knowledge status in statistics. These beliefs may also relate to their post experience of learning statistics and doing statistical tasks. Measures on students' expectancy in statistics may be task-specific or be very general. For example, task-specific measures asked students to estimate their ability on different statistics topics such as explaining probability value (*p*-value), interpreting the value of standard deviation, just to name a few (Finney & Schraw, 2003). These survey items were intended to measure specific self-ability estimations. In contrast, the general level of expectancy measures asked general questions to determine how well students would perform in the statistics classroom or how good they were at statistics (Schau et al., 1995). These survey items were supposed to measure the overall self-ability estimation on their success probability in their statistics class.

Previous studies have focused on the class-level expectancy on success in their statistics course or the task-level expectancy (i.e., success expectancy on specific topics). However, a class-level expectancy may be too general while a task-level expectancy may be too specific. My previous section argued for the necessity of mastering both conceptual and procedural knowledge in statistics. If a student could not recognize the requirement of mastering both types of knowledge, that student may hold an unbalanced expectancy. By "unbalanced," I meant students would put more emphasis on conceptual knowledge-based tasks and less emphasis on procedural knowledge-based tasks, or vice versa. Consequently, statistics students could have a high expectancy of succeeding in one type of task but a low expectancy of succeeding in another.

The subjective value that students have perceived for a given task is also important. In the early pioneering work on the expectancy-value model, it was argued that individuals varied in their level of incentive value on the accomplishment of a given task (Atkinson, 1957). The term "incentive" represents the relative attractiveness of a specific goal or unattractive of a consequence in a particular situation. A limitation of this definition has been raised with respect to its task-specific, laboratory-based settings. Successive researchers extended the original definition of "incentive" to a multi-dimensional construct including attainment value, interest value, utility value, and cost (Eccles et al., 1983). In particular, the first three dimensions (i.e., attainment, interest, and utility value) were empirically supported in subsequent confirmative factor analysis (Eccles & Wigfield, 1995). The dimension of cost was not directly related to my study and therefore it was excluded in my following review and discussion.

In the context of statistics, many researchers have studied students' subjective value of statistical knowledge, statistics course and/or learning of statistics. Three dimensions of the "value" component from the expectancy-value framework can be easily adapted in the statistics context. The attainment value on statistics will focus on the importance of doing well in a statistics task/course in relation to an individual's goals or personal life. The interest value will focus on the interest of a statistics learner in learning and applying statistical knowledge. The utility value will focus on the usefulness of the statistical knowledge/statistics course/the applications of statistics.

Here is an example of how those value dimensions can be operationalized in the field of statistics. If a student reports high in his/her attainment value on statistics, it means that person believes statistics is important to him/her (e.g., "statistics is important to me;" "doing well in statistics is meaningful to me"). If a student reports high in his/her interest value on statistics, it means the individual enjoys learning or applying statistical knowledge and likes to learn more about statistics. The enjoyment is intrinsically developed. Finally, if a student reports high in his/her utility value on statistics, it means this person believes statistics is valuable for his/her future career or life. For this reason, he/she will value the subject of statistics and the accomplishment of the statistics course. Empirical studies on these dimensions of statistics students' subjective value will be discussed in a later section.

Individuals' expectancy and their value will jointly determine how well they will perform in a task or performance scenario. Previous research has also focused on the relation between students' expectancy and value. Studies using different research designs have revealed a positive relation in general. Correlational studies found that there was a positive relation between expectancy and value (Eccles et al., 1983). Experimental studies (studies with interventions) indicated that there was a positive relation between expectancy and value, but it was significant only for students with low expectancy (Hulleman & Harackiewicz, 2009).

Theoretically, it is assumed that statistics students' expectancy and achievement value are positively correlated (Wigfield & Eccles, 1992). In terms of causal reference, several studies have demonstrated the expectancy component functions as a prior causal factor (Bandura, 1997). In my research, however, the causal relation regarding the expectancy and value components is not the focus of the study.

Finally, research found that both students' expectancy and their achievement value could positively predict their achievement in different academic disciplines. Expectancy has consistently functioned as one of the strongest psychological predictors of student achievement (Wigfield & Eccles, 2000). The effect of perceived value on achievement may vary from study to study (see Hulleman et al., 2010; Wigfield & Cambria, 2010; Wigfield, Tonks, & Klauda, 2016). One carefully designed study found that the value component did not predict achievement alone; however, the interaction between expectancy and value was a significant predictor for student achievement in the field of mathematics and English literacy (Trautwein et al., 2012).

Expectancy in Statistics

Statistics is often rated as one of the most difficult courses among non-statistics majored undergraduates across different programs (Garfield, 1995; Garfield & Ben-Zvi, 2007; Schau et al., 1995). Some researchers have put much efforts into reforming the course curriculum and improving students' learning strategies (Gal & Garfield, 1997; Finney & Schraw, 2003). In another approach, other researchers have focused on the motivation aspect such as students' expectancy for success in schools and classrooms.

As mentioned before, different researchers have developed different ways of conceptualizing students' expectancy for success in the given context. Although theoretical debates are still going on regarding what to include/exclude in the expectancy construct, I will focus on the empirical measurement tools which were developed by different researchers. In their original definition in the expectancy-value model, the expectancy component includes the expectation for future success as well as current self-ability beliefs (Wigfield &Eccles, 2000). To adopt this popular measure in the context of statistics, an individual learner's expectancy for success in statistics should be measured based on statistics students' expectations of how well they will do in their statistics class and their self-ability in doing statistics. This operationalized definition of individuals' expectancy should include ability-related beliefs about how well they do in specific contexts (in this case, their statistics class) and how well they will achieve in the future.

The measurement items could be easily adapted from the previous measurement scale into the context of the statistics class (Eccles & Wigfield, 1995). Some examples of adapted scale items could be: "Compared to other students, how well do you expect to do in statistics?" or "How good at statistics you are?" However, one possible drawback of adapting these survey items is that they are specified in the course level, and it may be too general to tap into some specific topics in the acquisition of statistical knowledge. In other words, students' self-reported course level expectancy may not be accurate enough to predict their actual performance in statistics tasks. In my study, those items were revised to make participants to reflect their expectancy in the context of conceptual and procedural statistical knowledge.

Schau and colleagues (1995) have developed a similar measurement scale regarding statistics students' expectancy for success in statistics classroom. This construct was labeled as cognitive competency; however, it is aligned with the definition of the expectancy component from the expectancy-value model (Schau & Emmioglu, 2012). The so-called "cognitive competency" survey scale included positively-worded items like "I can learn statistics" or "I will understand statistics equations" and negatively-worded items like "I will have troubles understanding statistics" or "I will make a lot of math errors in statistics". Similarly, one drawback of adapting this type of scale is that these questionnaire items still focus on the general context of statistics class. Statistics students tended to overestimate the probability of their success in the course performance if they reported on a too-general expectancy scale (Hall & Vence, 2010).

Some researchers in the statistics education field have developed survey items to measure task-specific, ability-related beliefs. For instance, undergraduate students' statistics self-efficacy was investigated in their statistics courses (Finney & Schraw, 2003; Olani, Hoekstra, Harskamp, & van der Werf, 2011). These survey items mainly focused on their self-ability ratings on understanding different statistical concepts in the intro statistics course. Examples of such survey items include "identify(ing) a distribution that is skewed" and "interpret(ing) the probability value (*p*-value) in a given statistics procedure." These studies conceptualized the statistics self-efficacy in the context of specific statistical topics. Nonetheless, those items were not constructed to reflect the future-oriented expectancy. In contrast, theory-based expectancy measures tended to include self-ability beliefs (Wigfield & Eccles, 2000).

In sum, previous studies on the expectancy component in statistics developed survey items that were either too general (e.g., Eccles & Wigfield, 1995; Schau et al., 1995) or too specific (e.g., Finney & Schraw, 2003; Olani et al., 2011). Survey items that are too general do not accurately reflect individual self-ability beliefs regarding specific topics in statistics. Survey items that are too specific do not reflect students' general understanding of the subject in their statistics class. In my study, I argue that measuring students' expectancy should be aligned with the level of the proposed statistical knowledge tests (e.g., expectancy for success in the conceptual knowledge test and expectancy for success in the procedural knowledge test). In addition, Eccles and Wigfield's scale (1995) was theory-based and therefore it reflected a top-down research approach. The scale developed by Schau et al. (1995) was designed for practical issues, and therefore it reflected a bottom-up research approach. In fact, a lot of their survey items overlapped with each other. I used Eccles and Wigfield's scale in my study design since I adopted the expectancy-value model.

Achievement Value on Statistics

Value, especially achievement value (subjective value), is the other focus in the expectancy-value model. Students' subjective value on tasks is crucial for determining their future behaviors or persistence (e.g., students may report that they will continue studying statistics since they highly value the learning subject). As mentioned before, three dimensions of students' subjective value—attainment value, interest value, and utility value contribute to the achievement value construct (Eccles & Wigfield, 1995; Wigfield & Eccles, 1992). Each of the value dimensions is discussed below.

Attainment value on statistics refers to the perceived importance of statistical knowledge and skills. In their original measurement scale, Eccles and Wigfield (1995) developed items such as “Is the amount of effort it will take to do well in ____ worthwhile to you?” and “Being good at solving problems which involves ____ is worthwhile.” Readers can simply fill in the blanks with the term “statistics” above. This dimension of the subjective value is supposed to describe the level of perceived importance of the given task in the given context and attainment value should influence the individual’s length of persistence on the task (Battle, 1965).

In the existing survey scales of measuring the students’ achievement value on statistics, none of them has explicitly addressed the dimension of attainment value. Instead, the “utility value” and “interest value” are often measured. Schau and colleagues (1995) developed one popular questionnaire and that questionnaire had one component “value” in the field of statistics. It was operationalized as “usefulness, relevance, and worth of statistics” (Schau et al., 1995, p. 870). Their definition was concentrated on the dimension of utility value. Similarly, this group of researchers had updated their questionnaire and added another value-related subscale “interest” (Schau, 2003). In sum, their up-to-date survey includes measurement items on both statistics students’ utility value and their interest value on statistics.

It is not a surprise that various scales labeled differently have evaluated the same construct in fact. Eccles and Wigfield (1995) developed the utility value scale which included items such as “How useful is learning ___ for what you want to do after you graduate and go to work” and “How useful is what you learned in ___ for your daily life outside school?” Similarly, Schau and colleagues’ survey items included negatively-worded items such as “Statistics is worthless" and "Statistics conclusions are rarely presented in everyday life" as well as positively-worded items such as "I use statistics in my everyday life" and "Statistical skills will make me more employable" (Dauphinee, Schau, & Stevens, 1997). The commonality of both scales has referred to the value of knowing and applying statistical knowledge and skills in learners' daily lives and in their careers. Statistics has utility value to students because they may think those knowledge and skills can help them achieve other goals (personal or future-related).

Regarding the dimension of interest value, Eccles and Wigfield (1995) developed the interest value scale to address interest or enjoyment when students work on a given task. The questions in the scale include items such as "In general, I find working on ___ assignments is very interesting" and "How much do you like doing ___?" Likewise, the "interest" scale developed from Schau and colleagues (Schau, 2003) is to measure self-reported individual interest in statistics. The questions in this scale include items like "I will enjoy taking statistics courses" and "I am interested in using statistics". The commonality of both scales is that they all refer to terms such as "enjoy/enjoyment" or "interested in/interesting". Statistics has an interest value to students because they may think doing it is enjoyable. Readers can easily figure out there are differences between utility value and interest value. In the literature, motivation researchers (Deci & Ryan, 1985; Wigfield & Eccles, 1992) also argue that individual interest value belongs to intrinsic motivation (intrinsic reasons for people's behavior) while individual utility value belongs to extrinsic motivation (extrinsic reasons for people's behavior).

Students' value on statistics may predict their future behavior and performance in their statistics classroom. A meta-analysis study revealed a small effect size on the relation between students' perceived value on statistics and their statistics achievement (Emmioglu & Capa-Aydin, 2012). Some researchers have argued that value is not a strong predictor of achievement

(Wigfield & Cambria, 2010). Instead, students' value on statistics may strongly predict their future behaviors or behavioral intentions (Wigfield & Eccles, 2000) such as continuing to learn statistics.

A group of researchers has conducted a series of motivation interventions focusing on students' value component (Hulleman et al., 2010; Hulleman, Kosovich, Barron, & Daniel, 2017). Evidence demonstrated that such interventions succeeded in improving undergraduate students' interest and utility value in different academic disciplines. Their studies aimed at manipulating students' utility value on a subject (e.g., how meaningful/useful is mathematics/psychology to their lives) by asking them to write an essay about the relevance of the subject which they investigated. Accordingly, I then proposed to adapt their approach but tailored it to the context of statistics.

In sum, students' subjective value on statistics is crucial for predicting their future behavior as well as their retention in statistics class. The dimensions of attainment value, utility value, and interest value portray a comprehensive picture of students' reasons for why they engage or disengage in studying statistical knowledge and skills. Existing measurement tools have overlapped to measure one or more than one dimension of the subjective value. As shown below, a motivation intervention focusing on learners' utility value has demonstrated its effectiveness to enhance individuals' perceived utility value and interest in different disciplines.

Motivation Intervention: Focusing on the Utility Value

Many motivation researchers in education field who aim at explaining students' motivation have tried to figure out the mechanism of how individuals' motivation influences their behavioral choice or their behavioral performance. In the framework of the expectancy-value model, the focus of interventions could be on the expectancy component or the value component. I choose to focus on the value components for two reasons. First, as mentioned before, the expectancy component was sometimes hard to change, especially in a short time period. Instead, researchers have figured out effective ways to facilitate the value component. The variance in perceived value could also contribute to expectancy change as well as performance improvement. Second, the knowledge responsibility prime which I proposed aimed at enhancing their beliefs about responsibility to use statistical knowledge in their statistics courses. It was a smooth transition from the statistical knowledge tests to remembering the value of statistics.

In a series of motivation intervention studies, Hulleman and colleagues (Hulleman & Harackiewicz, 2009; Hulleman et al., 2007, 2010; Rosenzweig et al., 2019) applied a utility value intervention approach to enhance students' perceptions of the relevance of the given discipline. For instance, high school students were assigned in a randomized experimental design in their science classes (Hulleman & Harackiewicz, 2009). The utility value intervention group received a writing intervention which aimed to reveal their utility value—the usefulness of the scientific knowledge and techniques in their lives or for future career. The post-test results showed the intervention group achieved higher grades and reported stronger continued interest in the science subjects compared to the control group. However, I did not make any prediction on participants' exam performance since my study material did not directly relate to any content or instruction in their introductory statistics course.

The effectiveness of the utility value intervention approach was reexamined recently among undergraduate students in the domain of mathematics, biology, and psychology (Hulleman & Barron, 2016; Hulleman et al., 2010). The mechanism of this intervention approach is supposed to help undergraduate students increase their perceptions of the usefulness of the learning subject and its relevance to their lives. It is also supposed to offer opportunities for students to make connections by themselves between the learning tasks they do and their daily lives—which in theory is a less threatening and more self-affirming type of intervention (Hulleman et al., 2010; Yeager & Walton, 2011). According to this research, such an indirect way of intervening could be more effective than overt strategies.

Based on the literature review, four possible factors may influence the effectiveness of this motivation intervention approach on students' motivational outcomes. The first factor is participants' expectancy level. The utility value intervention was more effective for those who reported lower in their expectancy for success in their performance (Hulleman et al., 2010). The second factor is the frequency (dosage) of the intervention. Hulleman and colleagues (2017) recently found that the more frequently undergraduate students saw the utility value on their chosen courses, the higher the performance they achieved in the classroom. The third factor is the time period of the post-intervention. The duration between the intervention session and the post-intervention/follow-up session varies from a couple of weeks to one academic semester (Acee & Weinstein, 2010; Hulleman et al., 2017). Finally, the timing of applying such interventions is crucial: the motivation intervention may be more effective especially at the beginning of the academic semester or school year (as suggested by Yeager & Walton, 2011).

I adapted the utility value intervention approach with some revisions. First, considering that the research activities were already complicated (especially for the group of reporting responsibility for knowledge and utility value), the one-time utility value task was adopted to reduce participants' fatigue. Second, unlike a previous study where the researchers introduced the utility value task in the second half of the semester (Hulleman et al., 2010), the utility value task was scheduled at the beginning of each semester based on the suggestion from Yeager and

Walton's review (2011). Finally, I didn't make any causal inference between value intervention and academic achievement by applying this approach. Instead, I asked undergraduate students to reflect on their experience in learning statistics and considered this was part of the responsibility reflection for predicting their commitment to learning statistics.

Commitment to Learning Statistics

Previous literature review has demonstrated different components that may affect individuals' responsibility beliefs about learning introductory statistics. The level of perceived responsibility may associate with the extent to which undergraduates report their commitment to learning statistics. Various constructs, such as learning on individuals' own demands and perseverance and focusing on the growth mindset and purposeful engagement all present specific indicators of their perceived commitment (Brown, 1988; Dweck, 2007). Here, commitment to learning statistics was used to represent a set of loosely defined motivational beliefs. Specific constructs were introduced below.

First, individuals' *persistence decisions* were explored. This construct is based on the intention to put in the effort and stick to it (Thorkildsen & Xing, 2016). Adding this variable was based on previous research (Schau & Emmioglu, 2012) where students' perception of investigating their "effort" in learning statistics declined over time. One problem in the previous study is that the perceived persistence should be considered along with the perceived ability in statistics.

Second, individuals' *perceived statistical ability* was proposed. This is an important indicator to reflect students' self-efficacy (Bandura, 1997). Perceived ability employed self-confidence and confidence by comparison with others. It was reported unchangeable in previous research among college students in their introductory statistics courses (Schau & Emmioglu, 2012). My study aimed to confirm whether perceived statistical ability was unchangeable (as consistent as in the previous research).

Third, it was crucial to investigate undergraduates' *attribution beliefs* in their statistics success and failure scenarios. Based on the literature (Weiner, 1979), I selected four common causes: ability, effort, difficulty of the task, and luck in the success and failure situations. Ability and effort are internal attribution causes while difficulty of the task and luck are external attribution causes. Internal attribution may represent forms of personal responsibility while external attribution may represent forms of excuses (Weiner, 1979). Whether individuals hold adaptive or maladaptive attribution beliefs may affect their future performance and emotions (Nicholls, 1989; Weiner, 1985).

Finally, the construct of general *attitude* toward statistics was adopted from previous research (Ajzen & Madden, 1986). This construct served as a generalized evaluation about statistics. It was dropped from later analysis since it did not help with explaining responsibility for learning statistics but offered a means of comparing the measurement properties of other, more informative variables. In Chapter 3 and 4, I still include the scale for measuring individuals' attitude toward statistics and descriptive results to showcase its vagueness.

The Current Study

By reviewing the previous research and extending other researchers' work, I designed my study as follows: (1) I designed to assign participants into one of the three study conditions: one group of participants completed conceptual and procedural statistical knowledge tests and questions about responsibility to use statistical knowledge (conceptual statistical knowledge prime and procedural statistical knowledge prime), as well as essays about relevance of statistics and questions that were about responsibility for perceived utility value of statistics and for remembering the utility value. This group was referred as the “full responsibility” group. A second group of participants completed conceptual statistical knowledge prime and procedural statistical knowledge prime only. This group was referred as the “knowledge responsibility” group. A third group of participants completed none of the knowledge primes and the utility value prime. Those participants only completed some statistics-irrelevant surveys. This group was referred as “undefined responsibility” group.

The design of responsibility groups was not an intervention for the introductory statistics course. Instead, undergraduate students who were assigned into the full responsibility and the knowledge responsibility group had the opportunities to explore their responsibility beliefs about using conceptual statistical knowledge and procedural statistical knowledge. Previous studies indicated that individuals may not have an equal understanding of each type of statistical knowledge. Therefore, it is essential to ask if undergraduates report equal responsibility for using those two types of statistical knowledge in the introductory statistics course.

In addition, undergraduates who were assigned into the full responsibility group reported their responsibility for perceiving the utility value of statistics. Undergraduate non-statistics majors may benefit from appreciating the relevance of statistics and reflecting their responsibility for perceiving the value of statistics. Also, this group of undergraduates had the opportunity to explore a bigger picture about their responsibility for learning statistics.

(2). Prior those research activities, all participants completed a questionnaire about their commitment to learning statistics. Then, they answered the similar questions in the middle and at the end of the semester. This design can help to explore both cross-sectional and longitudinal differences in individuals’ commitment to learning statistics.

(3). All participants also reported their demographic and academic background characteristics. Their previous learning experience was an important factor which may affect their commitment to learning statistics. The number of previous mathematics/statistics courses completed in their high schools is crucial to include as a covariate in the later analysis.

(4). At the end of the semester, I collected participants' statistics exam scores on their 1st quarter, midterm, and final exam although I did not assume my study would have an impact on their exam scores. Their aggregated exam score can serve as background information and can help to discuss the possible reasons for their commitment/retention in the introductory statistics course. Below, I have listed my research questions:

Research question 1. At baseline, will undergraduates who were randomly assigned in the full responsibility, the knowledge responsibility, or the undefined responsibility group report equally on their initial commitment to learning statistics? Will they report similar academic background information?

Research question 2. What is the relationship between conceptual statistical knowledge score and responsibility score of using conceptual statistical knowledge among the full and the knowledge responsibility group? What is the relationship between procedural statistical knowledge score and responsibility score of using procedural statistical knowledge among the full and the knowledge responsibility group?

Research question 3. Will undergraduates from the full responsibility, the knowledge responsibility, and the undefined responsibility group report differently on their commitment to learning statistics in the middle of the semester? Will their commitment to learning statistics differ after controlling for previous experience of learning mathematics/statistics?

Research question 4. Will undergraduates from the full responsibility, the knowledge responsibility, and the undefined responsibility group report differently on their commitment to learn statistics at the end of the semester? Will their commitment to learning statistics differ after controlling for previous experience of learning mathematics/statistics?

Research question 5. Regardless of the responsibility groups, will individual's commitment to learning statistics change over the semester? Will undergraduates from the full responsibility, the knowledge responsibility, and the undefined responsibility group differ in their statistics exam score?

Research question 6. Among the full responsibility group, will perceived responsibility beliefs from the knowledge primes and the utility value prime predict individuals' commitment to learning statistics in the middle and at the end of the semester?

Research question 7. Among the knowledge responsibility group, will perceived responsibility beliefs from the knowledge primes predict individuals' commitment to learning statistics in the middle and at the end of the semester?

Research question 8. What is the estimated retention rate in the introductory statistics course among all participants in this study?

CHAPTER 3

METHODS

The purpose of this study was to investigate three types of responsibility for learning statistics: responsibility to use rudimentary conceptual statistical knowledge, responsibility to use rudimentary procedural statistical knowledge, and responsibility for imagining the perceived utility value of statistics. A higher responsibility for learning statistics may foster a greater level of commitment to learning the subject matter. Participants recruited from introductory statistics courses at one Midwest public research university were randomly assigned to one of the three conditions. The "full responsibility" group completed responsibility primes that focused on their rudimentary conceptual and procedural statistical knowledge and their perceived utility value of statistics. The "knowledge responsibility" group completed knowledge responsibility primes that focused on rudimentary conceptual statistical knowledge and procedural statistical knowledge only. Finally, the "undefined responsibility" group completed activities that were irrelevant to statistics. All groups reported their commitment to learning statistics at the beginning (baseline), in the middle (1st post-test), and at the end (delayed post-test) of the semester. Individuals' statistics exam scores were collected at the end of the semester with the participants' permission.

Participants

Undergraduates enrolled in an introductory statistic course who volunteered to participate in the research were recruited. A total of 197 undergraduates were recruited from two courses taught by lecturers from the Statistics Department at one Midwest public research university. Thirty students either declined to participate later or never finished the initial study sessions. Finally, 167 participants (Median age = 19 years old) attended the research sessions at the beginning of the semester (response rate = 85%). Table I summarizes demographic information clustered by responsibility groups.

Table I

Demographic Information

	Full responsibility	Knowledge responsibility	Undefined responsibility	Total
Male	13	18	17	48
Female	39	37	42	118
African	5	3	2	10
Asian	14	14	23	51
Latino	13	17	16	46
Caucasian	9	12	12	33
Race/Ethnicity: Other	12	9	6	27

Note. Full responsibility: Responsibility for knowledge and motivation; Knowledge responsibility: Responsibility for knowledge only. Chi square statistics were conducted for detecting potential group differences in gender and in race/ethnicity. Results showed non-significant group difference was detected, $p > .05$. There was one who preferred not to say about the gender information in the full responsibility group.

Participants reported a wide range of majors and program levels. The summary of reported majors is presented in Table II. Ninety-five were freshmen, 41 were sophomores, 10 were juniors, five were seniors, and 16 reported "other".

Table II

Frequencies of Reported Majors

Major	Frequency (%)
Medicine/Pre-Medicine	40 (24.0)
Nursing/Pre-Nursing	33 (19.8)
Natural Sciences (e.g., Biology)	14 (8.4)
Social Sciences (e.g., Sociology)	12 (7.2)
(Applied) Health Sciences (e.g., Rehab Science)	13 (7.8)
Other majors	17 (10.2)
Not Decide/Don't know	38 (22.8)
Total	167

Note. The "Other" category included those who checked "other", who didn't check anything, or who reported majors which did not belong to those previous categories.

The targeted population was the undergraduates who were enrolled in introductory statistics courses because that population of statistics learners has increased in the past a few decades. For instance, about 508,000 students registered for an introductory statistics course in a two- or four-year college/university in 2010 across the United States (Carver et al., 2016). At the public research university where this study was done, students from a broad range of majors are required to take at least one introductory statistics course to finish their program. The participants in my study represents a convenience sample from the targeted population.

Procedures

This project was approved by the Institutional Review Board (IRB). At the beginning of each semester, I contacted statistics instructors. With their approval, I entered classrooms right before the lecture started and conducted a short, 5-minutes presentation about my research. Then, I distributed the hard copies of the consent forms to all students in the classroom and exited. I collected signed consent forms from the volunteers after class. To encourage their participation, a couple of days later, I visited the same classes and collected additional consent forms signed by the study volunteers. Using information provided on the consent forms, I scheduled appointments for meetings in an educational technology lab via email. Participants can reschedule their lab sessions one day before their scheduled sessions if they cannot make it.

Lab Sessions

Upon arrival, participants were assigned to use the lab computers for completing their lab sessions. They were randomly assigned to one of the three study conditions, but all completed paper-and-pencil activities and online survey activities. A list of research activities by responsibility group and their statistics exam schedules is shown in chronological order in Table III. In *paper-and-pencil* section, volunteers consented and reported their baseline commitment to learning statistics. The full responsibility group (i.e., taking responsibility for knowledge and perceived utility value) and knowledge responsibility group (i.e., taking responsibility for knowledge only) completed knowledge responsibility primes, which focused on rudimentary conceptual and procedural statistical knowledge as well as perceived responsibility to use such statistical knowledge. The undefined responsibility group completed non-statistics related surveys. In *online* section, the full responsibility group completed the utility value prime, which consisted of a short essay on the relevance of statistics to their lives (focusing on the responsibility for perceived utility value of statistics). They also reported their beliefs about responsibility for remembering the utility value of conceptual and procedural statistical knowledge. The other two responsibility groups completed non-statistics related surveys (general psychological surveys). At the end of the online session, volunteers reported their demographics and academic background information.

Table III

List of Research and Course Activities by Responsibility Group

Research and course activities	Full responsibility	Knowledge responsibility	Undefined responsibility
1 Consent (Appendix B)	x	x	x
2 Baseline commitment (Appendix C)	x	x	x
3 Conceptual statistical knowledge prime (Appendix D)	x	x	
4 Procedural statistical knowledge prime (Appendix D)	x	x	
5 Statistics utility value prime (Appendix E)	x		
6 Demographic/academic survey (Appendix F)	x	x	x
7 1st quarter statistics exam	x	x	x
8 1st post-test commitment (Appendix C)	x	x	x
9 Mid-term statistics exam	x	x	x
10 Delayed post-test commitment (Appendix C)	x	x	x
11 Final statistics exam	x	x	x

Note. Activities 1 – 3 were completed in paper-and-pencil format. Activities 4 – 6, 8, and 10 were completed online via Qualtrics. Activities 7, 9, and 11 were course exams. Study materials used in those research activities were presented in Appendix B – F.

Post-Test Surveys

One to three weeks after the lab sessions, volunteers received the 1^{st} post-test online survey. Every participant was asked to answer the same questions about their commitment to learning statistics in the current time framework. Surveys were distributed roughly after the 1^{st} quarter statistics exam and before the mid-term statistics exam. Towards the end of the semester, participants received another online survey (delayed post-test) to report their commitment to learning statistics across the semester. Individuals were also requested to offer their university email address as a linking variable to connect each set of survey answers with their lab session responses. Email reminders were sent out to politely encourage participants to complete the post-test surveys. I also stopped by at each of their classrooms to remind the participants about completing the post-test surveys.

Collecting statistics exam scores and debriefing. After the final statistic exam, I collected their 1^{st} quarter, mid-term, and final statistics exam scores. I provided statistics instructors a list of student names who signed the consent forms for giving the information about their exam scores. After those exam scores were matched with the participants' previous lab/questionnaire responses, I removed all the previous identifiers (i.e., university email address) from the dataset. Every participant in this study received a debriefing email that explained the purpose of the study with the principal investigator's contact information if any participant had any questions or concerns.

Managing attrition. At the time of participant recruitment, I asked all the participants to voluntarily offer their university email address. I used their emails to send them reminders for those who didn't have a chance to complete the research activities or who need to reschedule their sessions. In addition, I conducted a gift card lottery during one semester to encourage volunteers to complete the research activities. Participants who completed the lab sessions and at least one post-test survey were eligible for entering that lottery pool. I found there were no differences in any results attributable to this retention approach.

Materials

The measurement materials administered in this study included questionnaires/surveys, conceptual and procedural statistical knowledge tests, and an essay task about the relevance of statistics to students' daily lives. Instruments were introduced following the chronological order of research activities (See Table III).

Commitment to Learning Statistics

This set of questionnaires was completed as soon as students started the lab session and during post-test sessions and served as baseline measures. Participants were asked to complete the same questionnaire in the middle and at the end of the semester. This set of questionnaires consisted of several self-reported constructs of undergraduates' perceived commitment to learning statistics. First, persistence decisions were assessed using four items and each item was rated on a 7-point scale (1 – do not agree at all, 4 – neutral, 7 – totally agree). The persistence decision was focused on students' intention to making effort and investment in learning statistics. A sample item is: "I always try to complete my statistics homework and other assignments". Statistics of the sample size and internal consistency (Cronbach's α) for the baseline, 1st post-test, and delayed post-test are reported in Table IV. I adapted three items from a previous study (Thorkildsen & Xing, 2016) and then added one item "I figured continuing to learn statistics knowledge and skills is very important". All the items were revised to fit the context of learning statistics (See detail in Appendix C).

Second, participants' perceived statistical ability for completing their introductory statistics course was measured using three multiple choice items. The perceived statistics ability focused on the interpretation of their statistics ability in their class (self-rating and compared to others) and how well they expected to achieve in their statistics class. A sample item was "My ability in statistics class is…" Items were rated on a 7-point scale: 7 "excellent", 6 "very good", 5 "moderately good", 4 "average", 3 "moderately poor", 2 "poor", and 1 "very poor". Statistics of the sample size and internal consistency for the baseline, 1st post-test, and delayed post-test scale are reported in Table IV. This scale was adapted from Thorkildsen (2017) and those questions were tailored to the learners' perceived statistical ability (See Appendix C).

Table IV

Sample Size and Internal Consistency for Commitment to Learning Statistics

	Baseline *N* (α)	1st post-test *N* (α)	Delayed post-test *N* (α)
Persistence decisions (4 items)	167 (.82)	147 (.86)	96 (.80)
Perceived statistical ability (3 items)	167 (.86)	140 (.86)	98 (.88)
Success attributions (4 items)	167 (.53)	140 (.55)	98 (.59)
Failure attributions (4 items)	167 (.63)	140 (.56)	98 (.45)
Attitude toward statistics (1 item)	160	119	98

Note. For success and failure attribution, Cronbach's α was calculated although the item level statistics were used for later analysis. Attitude toward statistics was not a scale score so Cronbach's α was not applicable to report.

Third, attribution beliefs in statistics was measured in two scenarios: experiencing success and failure in the introductory statistics course. This questionnaire assessed students' perceived *causes* of their academic achievement or failure in their statistics class. In the success attribution scenario, students were asked to rate four items representing their possible reasons for their statistics success: ability, effort, easy task, and luckiness (based on a 7-point scale; 7 – strongly agree, 4 – neutral, 1 – strongly disagree). In the failure attribution scenario, students were asked to rate another four items representing their possible reasons for their statistics failure: low ability, little effort, difficult task, and bad luck. Statistics of the sample size and internal consistency for the baseline, 1st post-test, and delayed post-test scale are reported in Table IV. Questionnaire's items were adapted from Weiner (1985) and those items and scenarios were tailored to the context of statistics class. Each of those items in the success/failure scenario represents a unique dimension of individuals' attribution-the analysis in Chapter 4 on attribution beliefs was conducted at the item level.

Finally, students' general attitude about statistics was measured by one item: "On a scale of 0 – 100, how do you like statistics?". A possible number of their responses was between 0 and 100 with the integral value (e.g., 70, 85, etc.). Sample sizes for the baseline, 1st post-test, and delayed post-test item are reported in Table IV. This attitude measure was created for the context of learning statistics and reflected a holistic attitude about whether a learner had a favorable or unfavorable evaluation of statistics (Ajzen & Madden, 1986).

Rudimentary Statistical Knowledge Primes

Participants assigned in the full responsibility group and the knowledge responsibility group both completed their rudimentary statistical knowledge primes. In their statistical knowledge prime, they completed two sets of rudimentary statistical knowledge tests: One test focused on rudimentary conceptual statistical knowledge and the other focused on rudimentary procedural statistical knowledge. In addition, participants also answered the questionnaire of their beliefs about responsibility to use conceptual/procedural statistical knowledge before and after each knowledge test was administrated. In the next sections, (1) responsibility questionnaires, (2) rudimentary conceptual statistical knowledge test, and (3) rudimentary procedural statistical knowledge test were introduced.

Responsibility to use conceptual/procedural statistical knowledge. Within each statistical knowledge test, a K-W-L framework (Ogle, 1986) was adapted to evaluate the effect of the participants' perceived beliefs of their responsibility to use the type of statistical knowledge. The "know" step is to let students recall what they know regarding a specific learning task. This step helps students create their own reason for the learning task. The "want (to know)" step focuses on students' self-designed learning goals for the specific task. Finally, the "learned" step focuses on what students have learned or if the task was successfully mastered. This step creates an opportunity to consolidate their learning.

The K-W-L framework was adapted in my study. The K, W, and L steps were conducted before the knowledge test happens, when the test happens, and after the test happens. Specifically, before participants worked on their rudimentary statistical knowledge tests, they answered questions about the expectation for their performance on the knowledge test–the percentage which they expected to have for correct responses (range: 0 – 100%) and the importance of the statistical knowledge in those knowledge tests (on a 5-point scale; 1: not at all, 2: slightly, 3: moderate, 4: very, and 5: extremely). After completing the knowledge tests and reading the feedback, undergraduate students responded to the same questions, but the context of those questions emphasized the expectation for their performance on such knowledge tests and importance of similar knowledge for *the future*.

A sample of the pre-test item asking about the expectation for performance on the knowledge test is "Approximately, what percentage of correct responses you will expect to earn in the task? Circle one point from the scale of 0% - 100%." A sample of the pre-test item asking about the importance of the type of statistical knowledge before the test is "Do you think the ____(conceptual/procedural) statistical knowledge important to learn?". The post-test items were very similar as those pre-test items. The sample size for each item divided by the responsibility groups is reported in Table V. The complete list of the questionnaire items is shown in Appendix D, under the section of "perceived responsibility for conceptual statistical knowledge" and "perceived responsibility for procedural statistical knowledge".

Table V

Sample Size and Internal Consistency for the Conceptual and Procedural Statistical Knowledge Prime

Domain	Scales/Question Items	Full *N* (α)	Knowledge *N* (α)
Conceptual Statistical Knowledge prime	Pre-test expected performance (1 item)	51	48
	Pre-test importance (1 item)	45	48
	Post-test expected performance (1 item)	45	42
	Post-test importance (1 item)	45	42
	Conceptual statistical knowledge score (14 items)	45 (.81)	46 (.83)
Procedural Statistical Knowledge prime	Pre-test importance (1 item)	40	37
	Pre-test expected performance (1 item)	40	37
	Post-test importance (1 item)	40	37
	Post-test expected performance (1 item)	40	37
	Procedural statistical knowledge score (5 items)	40 (.76)	37 (.83)

Note. Full: Full responsibility group; Knowledge: Knowledge responsibility group. For individual questions instead of scales, Cronbach's α was not applicable and therefore was not reported.

Rudimentary conceptual statistical knowledge test. A statistics topic was selected to assess students' conceptual knowledge of "probability" and some related concepts (Cobb, 1992). Probability is a basic concept in statistics and undergraduates should already understand the basic idea about probability in their high school (Mathematics Standards, 2019). Concept map was an appropriate assessment tool to assess conceptual knowledge (Novak, 1990). In this study, the instrument was a concept-map-format test. The concept map was filled with related concepts and links that were connected to the key concept "probability". A sample item is "Probability: one common pattern follows ____", and the correct concept for this is "normal distribution".

Starting with a map designed by Cravalho (2010), the initial concept map task was reviewed by two doctoral students in statistics department who were teaching assistants in the introductory statistics course to make sure the content was appropriate for testing rudimentary conceptual statistical knowledge. Based on their review suggestions, the final concept map task included 14 select-and-match type of questions. To simplify the test, the higher-order concept "probability" was given to students as a known concept. Students were then asked to match supplied concepts/links with their position in the map. There were 7 options for concepts: (A) Comparing scores (z-scores) from different normal distributions, (B) each individual in the population has an equal chance of being selected, (C) no bias, (D) proportion, (E) random sampling, (F) the normal distribution, and (G) the observed probability of any particular outcome that can happen when several different outcomes are possible. And there were 7 options for links: (H) can be described by, (I) fulfill the condition of, (J) insuring, (K) can be observed as, (L) ranges, (M) can be converted, and (N) is for.

Each correct match was scored 1, and the highest possible score was 14 while the lowest possible score was 0. Based on the scoring answer sheet (see details in Appendix D), two independent raters (a graduate student and me) scored about 15% of their responses. The inter-rater agreement on those scores is 100%. I then scored the rest of their responses. Statistics of the sample size and internal consistency of the conceptual statistical knowledge score are reported in Table V. After completing the concept map test, students were asked to review the standardized answer sheet and compare their answers to the standard answers.

Rudimentary procedural statistical knowledge test. Participants completed a procedural statistical knowledge test which focused on statistics learners' ability to recognize and use algebraic/arithmetic symbols, complete appropriate calculations, and solve statistical (e.g., probability) problems. The procedural statistical knowledge was a predictor for individuals' success in the introductory statistics course (Galli et al., 2011). Participants were asked to complete five constructed response type items: two items focused on the algebraic/arithmetic calculations and other three items focused on solving probability problems (hints and formulas were given to participants). An item example is "Solve the following expression: $(17\text{-}13)^2 + (10\text{-}13)^2 + (13\text{-}13)^2 + (12\text{-}13)^2$". Participants were asked to write down their answer as well as the steps for solving those problems.

Each item was scored as 2, 1, or 0 depending on correctness of the procedures that participants presented in their step-by-step solutions. A score of "2" indicated that participants correctly used the procedures to get the right answer to the given task; A score of "1" (partial credit) indicated that participants were *partially* correct about using right procedures; and a score of "0" indicated that participants did not use a correct solution procedure (Additional details for the scoring rubrics are in Appendix D, Table A2). A composite score was created by adding up those 5 item scores (the composite score range: 0 – 10). Statistics of the sample size and internal consistency of the conceptual statistical knowledge score are reported in Table V.

The content of the procedural statistical knowledge test was adapted from previous research (Galli et al., 2011; Schutz et al., 1998), and included arithmetic and algebraic skills (arithmetic calculation, equation, and statistics symbol/calculation; question #1, #3, and #4) and the knowledge of carrying out statistics procedures (independent probability and conditional probability; question #2 and #5). Two doctoral students in the statistics department who were teaching assistants in the introductory statistics course reviewed and confirmed the content of the test was appropriate for testing procedural statistical knowledge. Based on the scoring rubrics, two raters (a graduate student and me) scored about 15% of the participants' responses. The first-round rater agreement was 75%. The second-round rater agreement was 100%. After the second round, I scored the rest of their item responses.

The Utility Value Prime

Participants assigned to the full responsibility group completed a set of online instruments for imagining the relevance of statistics to their daily lives and reporting their responsibilities for perceiving the utility value of statistics. There were four instruments that the full responsibility group completed: (1) an essay task of reflecting the relevance of statistics to their lives, (2) a survey on their beliefs about the responsibility for perceived utility value of statistics, and (3) two questionnaires on their responsibilities for remembering the conceptual statistical knowledge and the procedural statistical knowledge.

Relevance of statistics essays. Participants were assigned to write one short essay by imaging the relevance of statistics to their daily lives. The essay-writing task aimed at providing an opportunity to motivate participants to think of the usefulness of statistics and to connect it to their personal use. Participants were asked to type their essays (1 – 3 paragraphs) focusing on explaining the potential usefulness of statistics to their daily life and giving 1 – 2 examples using their assigned computers in the lab. Essays were scored based on the number of personal pronouns (e.g., I, me, mine, us, our, ours) which indicted the extent to which participants made connections between statistics and their daily lives. This is based on previous study that the number of mentioning personal pronouns represented to what extent the respondents related the subject to themselves (Hulleman et al., 2010). The number of examples about relevance of statistics was also collected. A composite score was calculated by adding the number of personal pronouns and the number of examples. The sample size is reported in Table VI.

This task was adapted from previous research on utility value motivation intervention study (Hulleman et al., 2010) and two changes were made. First, the original task was to write a letter about usefulness of psychology to a significant person in participants' life or to social media. The task in my study was changed to be more personal, asking participants to connect their imagined relevance of statistics to their personal use and the context was changed to the usefulness of statistics. It is assumed that imaging the utility value of statistics to personal use may raise participants' awareness of their perceived utility value of statistics. In addition, the time for assigning the task to participants in the previous study was at the second half of the semester. In my study, it was assigned to participants at the beginning of the semester (suggested by Yeager & Walton, 2011) so they could get their earlier exposure on reflecting their utility value of statistics.

Table VI

Sample Size and Internal Consistency for the Utility Value of Statistics Questionnaire

Scales/Question Items	*N* (α)
Number of personal pronouns and examples	40
Perceived responsibility for statistics relevance (4 items)	40 (.88)
Remembering the value of conceptual statistical knowledge (12 items)	38 (.89)
Remembering the value of procedural statistical knowledge (12 items)	38 (.92)

Note. The number of personal pronouns and examples was not a scale score; therefore, reporting Cronbach's α was not applicable.

Responsibility for perceived utility value. Previous study did not include a measure of responsibility for perceived utility value (Hulleman et al., 2010). I created 4 questions asking participants to rate their beliefs about perceived utility value of statistics after the completion of the essay assignment on the relevance of statistics. Participants used a 5-point scale to evaluate: (1) How relevant is statistics to undergraduates in general? (2) how is statistics relevant to you? (3) how important is statistics to undergraduates' daily life in general? and (4) how important is statistics to your daily life? The 5-point scale response options are: 1 "not at all", 2 "slightly", 3 "moderately", 4 "very", and 5 "extremely". An average score was created to reflect participants' beliefs about responsibility for perceiving utility value of statistics (the higher their perceived responsibility, the more relevant participants feel about the statistics knowledge to their personal use).

Responsibility for remembering the utility value of statistics. I adopted the items from the measures of expectancy-value model in the previous literature (Eccles & Wigfield, 1995) and changed the subject to the utility value of conceptual and procedural statistical knowledge. Those items were administrated to participants who were assigned to the full responsibility group after they completed their essays and the survey about their responsibility for perceived utility value of statistics. This instrument was assumed to raise their responsibility awareness for remembering the utility value of statistics. It was administered twice: one in the context of conceptual statistical knowledge and another in the context of procedural statistical knowledge.

Initially, five items measured to what extent participants expected to success in mastering conceptual/procedural statistical knowledge. Seven items measured how much value participants put in conceptual/procedural statistical knowledge. Items were rated based on the 7-point scale (1: totally disagree; 4: neutral; 7: totally agree). For the utility value of conceptual and procedural statistical knowledge, the expectancy and value component were highly correlated in my study, $r = .86, .79$, $p < .001$, respectively. Therefore, a scale score was created to represent undergraduates' responsibility for remembering the utility value of conceptual statistical knowledge and for remembering the utility value of procedural statistical knowledge. Statistics of the sample size and internal consistency for those two scales are reported in Table VI.

Demographic and Academic Background

Every participant completed a demographic and academic background survey at the end of their lab session. The demographic questions asked the information about participants' age, gender, and race and ethnicity. The academic background survey questions asked the information about credits earned at the university, number of mathematics/statistics courses completed in high school, number of completed math/statistics course at college, a checklist of high school math/statistics courses, current (estimated) GPA, and academic majors. Among those questions, the survey question about the number of mathematics/statistics courses completed in high school was adopted from previous study (Cashin & Elmore, 2005; Schau et al., 1995). This is reported as an important factor in previous studies, and I use it as a covariate in my later analysis. Then, participants reported specific mathematics/statistics courses they completed from a list of mathematics/statistics courses offered in American high schools (Mathematics education in the United States, n.d.).

Statistics Exam

The 1st quarter, mid-term, and final exam scores in participants' introductory statistics course were collected and used as one important background variable for exploring the undergraduate students' commitment to learning statistics. Each exam included multiple-choice items and write-in items and the range of the possible row score for each exam was 0 – 100. The descriptive statistics of participants' statistics exam scores were reported in the Chapter 4. A *z* score and a *T score* were created to minimize the effect of test difficulty on the exam score on different occasions.

Data Analysis Plan

This study included lab activities which were administrated at the beginning of the semester and post-test surveys that were administered in the middle and at the end of the semester. Over time, there were sample attritions. Specific valid sample sizes for different measures across the semester can be found in Table IV – VI. Data analysis was conducted using

SPSS 25.0 (IBM Corp, 2017). I first examined the baseline measures and those measures showed adequate variance and reasonable psychometric property. I also compared scores on different beliefs and attributions to identify the best means of depicting students' commitment to learning statistics. I then compared measures of students' baseline commitment to learning statistics and of their responsibility primes across groups to ensure the group equivalence prior to introducing the responsibility primes. Those results helped to answer the research question 1.

Participants from the full responsibility and the knowledge responsibility group both completed conceptual statistical knowledge prime and procedural statistical knowledge prime. The relationship between their conceptual/procedural statistical knowledge score and their perceived responsibility score was examined using correlational analysis. The results from correlational analyses could answer the research question 2.

Focusing on undergraduates' 1st post-test commitment to learning statistics (in the middle of the semester), since different commitment components were supposed to correlated (for example, persistence decision was correlated with perceived statistical ability), MANOVA analysis was conducted for overall commitment scores (persistence decision and perceived statistical ability) and specific attribution commitment scores (different factors that individuals attribute their statistics success/failure to) while using responsibility groups (the full, the knowledge, and the undefined responsibility group) as an independent variable. Additionally, MANCOVA was conducted by adding individuals' previous mathematics/statistics experience as a covariate to confirm if any results from MANOVA stay the same. Those analysis procedures helped to answer the research question 3.

Similarly, for undergraduates' delayed post-test commitment to learning statistics (at the end of the semester), MANOVA analysis was conducted for overall commitment scores (persistence decision and perceived statistical ability) and specific attribution commitment scores (different factors that individuals attribute their statistics success/failure to) while using responsibility groups (the full, the knowledge, and the undefined responsibility group) as an independent variable. Additionally, MANCOVA was conducted by adding individuals' previous mathematics/statistics experience as a covariate to confirm if any results from MANOVA will remain the same. Those analysis procedures helped to answer the research question 4.

I use the following procedures to review the results of MANOVA analyses: First, I use the multivariate test results to determine if there is a main effect from responsibility group. I would report post hoc analysis results (with Bonferroni corrections) if there is any significant main effect. Second, I report results from between-subject tests if there is any significant effect on any single dependent commitment score. For any significant finding from those ominous tests, the observed power was reported. I use similar procedures to review the results of MANCOVA analysis. Box's test of equality of covariance matrices and Levene's test of equality of error variances were conducted. Those analysis results were not reported in Chapter 4 unless there was any significant inequality.

Comparisons of participants' commitment to learning statistics between the baseline and their two post-tests revealed a marked attrition rate. Although the fully saturated model will not be tested, I conducted the within-subjects ANOVA on individuals' persistence decisions, perceived statistical ability, beliefs about success caused by ability/effort/easy task, and beliefs about failure caused by low ability/little effort/difficult task which were reported 3 times: at the beginning, in the middle, and at the end of the semester. I examined if there were changes on those commitment constructs across the semester regardless of the responsibility groups (the membership of their responsibility groups was treated as the between-subject factor and was partialed-out during the analysis). I also compared individuals' aggregated statistics exam score by responsibility group. Those analysis results helped to answer the research question 5.

To answer the research question 6 and 7, I conducted the regression analysis in the full responsibility group to test if participants' beliefs about their responsibility for learning statistics would predict their persistence decisions and perceived statistical ability. I repeated the similar analysis for participants in the knowledge responsibility group using beliefs about their responsibility for learning statistics. Significant predictors were reported, and multiple tests were run for examining if the requirements for assuming linear regressions were met. The analysis results for testing linear regression assumptions were not reported in Chapter 4 unless there was any significant violation of those assumptions.

Finally, the information about participants' statistics exam scores in the introductory statistics course collected from those who consented to record their statistics exam scores can be an indicator for estimating the retention in their statistics course (research question 8). Descriptive statistics on their statistics exam scores were also reported—analysis results showed the content of the course tended to be harder for the participants from the beginning to the end of the semester.

CHAPTER 4

RESULTS

All participants (N = 167) who attended the research sessions were included in the final analysis. The focus of the study is to investigate undergraduate students' understanding of 3 types of responsibility (responsibility for learning conceptual statistical knowledge, responsibility for learning procedural statistical knowledge, and responsibility for perceiving the utility value of statistics) and their commitment to learning statistics in their introductory statistics course. This chapter was structured to answer the following research questions (the complete list of research questions can be found in Chapter 2):

Questions about the baseline: Will participants who were assigned in different responsibility groups report equivalently on the initial commitment to learning statistics as well as academic background? What is the relationship between conceptual/procedural statistical knowledge score and responsibility score of using conceptual/procedural statistical knowledge?

Question about their commitment to learning statistics: Will participants who were assigned in different responsibility groups report differently on their commitment to learning statistics in the middle and at the end of the semester after controlling the previous experience?

Question about changes of commitment to learning statistics: Will individual's commitment to learning statistics change across the semester regardless of their responsibility groups?

Question about the within-group responsibility variance: Will perceived responsibility belief from the knowledge primes (and the utility value prime) predict individuals' commitment to learning statistics for the knowledge responsibility group (and the full responsibility group)?

Question about retention: What is the estimated retention rate in the introductory statistics course among all participants in this study?

Variability in Baseline Commitment to Learning Statistics and Academic Background

Before testing the primary predictions, it is important to ensure that scores across participants show an adequate level of variance. This was established by first checking the relative equivalence of baseline commitment scores across the 3 responsibility groups and distributions of academic background characteristics.

Baseline Commitment to Learning Statistics

Descriptive statistics of undergraduates' perceptions of their (1) persistence decisions, (2) statistical ability, (3) success and failure attributions in statistics performance, and (4) general attitude about statistics are reported in Table VII.

Table VII

Descriptive Statistics of Baseline Commitment to Learning Statistics

Commitment to learning statistics	Full responsibility (n = 53)		Knowledge responsibility (n = 55)		Undefined responsibility (n = 59)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Persistence decisions	5.31	1.29	5.45	1.07	5.01	1.04
Perceived statistical ability	4.99	0.93	4.93	1.09	4.64	1.15
Success caused by						
Ability	4.43	1.56	4.67	1.48	4.40	1.43
Effort	5.60	1.17	5.84	1.10	5.23	1.35
Easy task	4.38	1.42	4.67	1.35	4.45	1.17
Good luck	3.73	1.67	3.57	1.65	3.96	1.52
Failure caused by						
Low ability	3.49	1.66	3.73	1.85	3.72	1.54
Little effort	4.36	1.78	4.05	2.17	3.71	1.73
Difficult task	4.60	1.38	4.20	1.56	4.34	1.55
Bad luck	3.19	2.01	2.78	1.66	3.48	1.72
Attitude about statistics	64.37	24.38	70.38	21.14	64.36	19.07

Note. Persistence decisions, perceived statistical ability, and attribution beliefs were based on a 1-7 scale; The general attitude about statistics was based on a 0-100 scale.

In addition, the normality of baseline commitment measures was checked. Skewness and kurtosis statistics are reported on baseline commitment in Table VIII. I cited a widely used criterion (Chou & Bentler, 1995) that a skewness value within $\pm$ 2 with a 95% confidence interval (CI) or a kurtosis value within $\pm$ 3 with a 95% CI was categorized as no severe violation of normality. Results in Table VIII demonstrated that all baseline component scores did not violate the assumption of normality.

Table VIII

Normality of Baseline Commitment to Learning Statistics

	Skewness (95% CI)	Kurtosis (95% CI)
Persistence decisions	–.88 (–1.25, –.51)	1.18 (.45, 1.91)
Perceived statistical ability	–.39 (–.76, –.02)	.35 (–.39, 1.08)
Success caused by		
Ability	–.42 (–.79, –.05)	.12 (–.61, .86)
Effort	–.76 (–1.13, –.40)	.69 (–.05, 1.42)
Easy task	–.33 (–.70, .04)	.31 (–.42, 1.04)
Good luck	.06 (–.31, .43)	–.58 (–1.32, .16)
Failure caused by		
Low ability	.23 (–.13, .60)	–.56 (–1.29, .17)
Little effort	–.01 (–.38, .36)	–1.04 (–1.78, –.31)
Difficult task	–.48 (–.85, –.11)	–.22 (–.95, .52)
Bad luck	.40 (.03, .76)	–.91 (–1.64, –.18)
Attitude about statistics	–1.14 (–1.52, –.76)	.92 (.17, 1.67)

Research question 1. At baseline, will undergraduates who were randomly assigned in the full responsibility, the knowledge responsibility, and the undefined responsibility group report equally on their initial commitment to learning statistics and similarly in their academic background?

Answer to research question 1. One-way ANOVA was conducted to compare whether these commitments to learning statistics scores differed by responsibility groups. Analysis results confirmed equal variances of scores across responsibility groups on baseline measures of persistence decisions, perceived statistical ability, success caused by ability/effort/easy task/luck, and failure caused by low ability/little effort/difficult task/bad luck, and attitude toward statistics, $p > .05$ after Bonferroni's correction for the number of tests was applied.

Dropping attitude toward statistics. Descriptive statistics of attitude toward statistics showed the highest skewness (Table VIII). Distribution scores for participants' attitude toward statistics are depicted in Figure 2. The attitude scores were too general and positively skewed, and they were not sensitive enough to detect group differences in specific aspects of students' commitments to learning statistics. This attitude variable was dropped from later analysis.

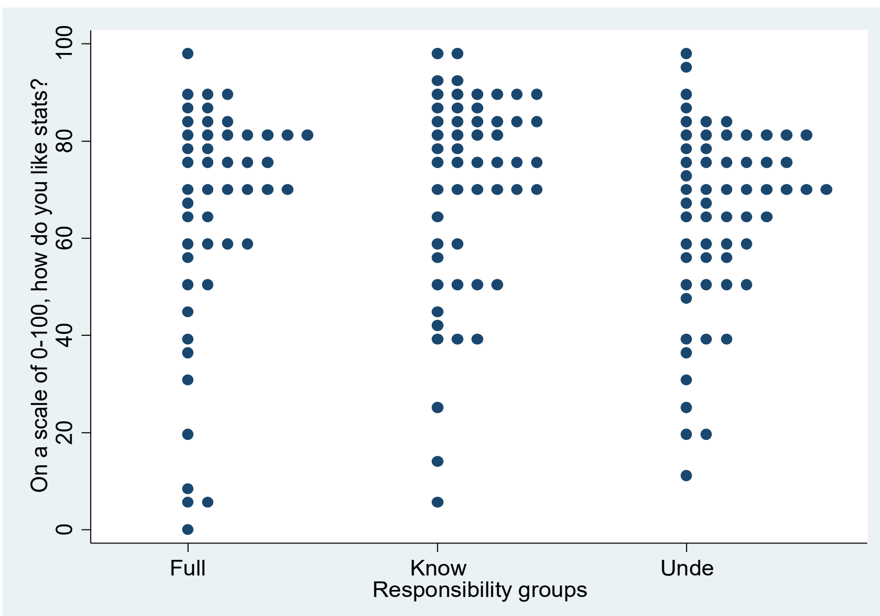

Figure 2. *Dotplot of participants' attitude toward statistics by responsibility groups.*

Academic Background. Descriptive statistics for academic background characteristics are summarized in Table IX. Results indicated that participants across the 3 responsibility groups reported equivalent academic background characteristics.

Table IX

Descriptive Statistics of Participants' Academic Background

Academic background	Full responsibility		Knowledge responsibility		Undefined responsibility	
	N	*M (SD)*	*N*	*M (SD)*	*N*	*M (SD)*
GPA	35	3.34 (0.39)	41	3.17 (0.55)	52	3.28 (0.45)
# of HS math	40	3.90 (1.22)	39	3.72 (1.52)	58	3.69 (1.47)
# of U math	38	0.37 (0.68)	37	0.86 (1.13)	58	1.02 (1.18)

Note**.** **GPA**: Grade Point Average. **# of HS math**: Numbers of completed high school mathematics/statistics courses. **# of U math**: Numbers of completed university-level mathematics/statistics courses. ANOVA analysis between responsibility groups did not detect any significant group difference in GPA, number of high school mathematics courses, and number of university mathematics courses, $p > .05$.

A closer look at the specific mathematics/statistics courses that participants have completed in high school also show comparable percentages across groups (Table X). Participants across the 3 responsibility groups reported similar percentages of completed high school mathematics/statistics courses, but more students completed algebra, geometry, and pre-calculus courses than calculus, trigonometry, or statistics courses. Results indicated this group of participants had the adequate level of mathematics/statistics background for learning introductory statistics.

Table X

Percentages of Completed High School Mathematics/Statistics Courses

Names of courses	Full responsibility (n = 43)	Knowledge responsibility (n = 40)	Undefined responsibility (n = 58)
Algebra I	74.4	75.0	70.7
Geometry	83.7	87.5	84.5
Algebra II	72.1	72.5	62.1
Pre-Calculus	67.4	82.5	69.0
Calculus	27.9	40.0	37.9
Trigonometry	55.8	57.5	48.3
Statistics	14.0	22.5	22.4

Note**.** Kruskal-Wallis test (an alternative, non-parametric equivalent of the one-way ANOVA) was conducted to compare the effect from responsibility groups on their percentages of completed mathematics and statistics courses in their high school. Nonsignificant group difference was detected, $p > .05$.

To summarize the statistical results from Table IX and X, participants' academic ability and educational background in mathematics/statistics was confirmed equivalent across the 3 responsibility groups. Previous research found learners' previous academic or mathematics/statistics experience impacted their motivation to learn statistics. The results here supported that it was safe to proceed further analysis when making conclusions about whether the responsibility primes would be useful reminders of participants' responsibilities for learning statistics when individuals reported similar academic background characteristics.

Variability in Conceptual and Procedural Statistical Knowledge Prime

The knowledge primes included two rudimentary statistical knowledge tests (conceptual and procedural) as well as participants' beliefs about their responsibility for connecting existed statistical knowledge to new knowledge in introductory statistics courses. Before testing the research predictions, descriptive statistics were compared to ensure an adequate level of variability in each set of the knowledge prime instrument.

Variability in Conceptual and Procedural Statistical Knowledge Test

Descriptive statistics of *conceptual statistical knowledge* score for participants in the full responsibility group (n = 45) and the knowledge responsibility group (n = 46) were reported, M = 4.58 (SD = 3.08), and M = 4.48 (SD = 3.29). The skewness for participants in the full and the knowledge responsibility group was .30 and .31. The kurtosis for participants in the full and knowledge responsibility group was –.52 and –1.09. Their scores were slightly positively skewed, which indicated there were more participants tended to score lower than the mean score. The result of negative kurtosis meant their test score distributions were flatter comparing to the normal distribution, indicating that the test scores were more spread out instead of grouping together near the mean. This is understandable since the majority participants scored lower than the middle point of the possible total score and no one scored higher than 11.

Descriptive statistics of *procedural statistical knowledge* score for participants in the full (n = 40) and the knowledge-only responsibility group (n = 37) were reported, M = 2.84 (SD = 3.07), M = 2.71 (SD = 3.27). The skewness for participants in the full and knowledge responsibility group was .71. and .90; the kurtosis for participants in the full and knowledge responsibility group was –.70 and –.21. Their scores were positively skewed, which indicated that more participants tended to score lower than their mean score; the negative kurtosis meant their test score distributions were flatter comparing to the normal distribution, indicating that scores were likely to spread out from the group mean.

Relationship between Statistical Knowledge Scores and Responsibility Scores

Research question 2. What is the relationship between conceptual statistical knowledge score and responsibility score of using conceptual statistical knowledge among the full and the knowledge responsibility group? What is the relationship between procedural statistical knowledge score and responsibility score of using procedural statistical knowledge among the full and the knowledge responsibility group?

Answer to research question 2. The answer to this question was two-fold. First, results of two-tailed correlational analyses on conceptual statistical knowledge, expectation for performance on conceptual statistical knowledge, and importance of conceptual statistical knowledge are presented in Table XI. Participants who reported stronger beliefs about the expectation for performance on conceptual statistical knowledge test tended to report stronger beliefs about the importance of conceptual statistical knowledge and higher test scores, $p < .05$, regardless of their responsibility groups. However, there was no significant relationship between individuals' conceptual statistical knowledge score and their beliefs about the importance of conceptual statistical knowledge among the full responsibility group; there was a significant positive relationship between individuals' conceptual statistical knowledge score and their beliefs about the expectation of performance on conceptual statistical knowledge test, $p < .05$.

Table XI

Correlations between Conceptual Statistical Knowledge and Responsibility to Use Conceptual Statistical Knowledge

Pearson's correlation	Full responsibility			Knowledge responsibility		
	Conceptual score	Performance	Importance	Conceptual score	Performance	Importance
Conceptual score	--			--		
Performance	.35* (n = 45)	--		.52** (n = 41)	--	
Importance	.18 (n = 45)	.36* (n = 43)	--	.33** (n = 39)	.37* (n = 39)	--

Note. **Conceptual score**: Conceptual statistical knowledge test score. **Performance**: Post-test expectation for performance on conceptual statistical knowledge test. **Importance**: Post-test importance of conceptual statistical knowledge. Significance level: *, $p < .05$, **, $p < .01$, ***, $p < .001$.

Second, two-tailed correlational analysis on procedural statistical knowledge, expectation for performance on procedural statistical knowledge, and importance of procedural statistical knowledge are presented in Table XII. Participants who reported stronger beliefs about the expectation for performance tended to score higher in procedural statistical knowledge test and report stronger beliefs about the importance of procedural knowledge, $p < .05$ regardless of responsibility groups. However, there was no relationship between individuals' procedural statistical knowledge score and their beliefs about the importance of procedural statistical knowledge among the full responsibility group. There was a positive relationship between procedural statistical knowledge score and the importance of procedural statistical knowledge among the knowledge responsibility group, $p <.05$.

Table XII

Correlations between Procedural Statistical Knowledge and Responsibility to Use Procedural Statistical Knowledge

Pearson's correlation	Full responsibility			Knowledge responsibility		
	Procedural score	Performance	Importance	Procedural score	Performance	Importance
Procedural score	--			--		
Performance	.59*** (n = 40)	--		.71*** (n = 37)	--	
Importance	.22 (n = 40)	.33* (n = 40)	--	.52** (n = 37)	.49** (n = 37)	--

Note. **Procedural score**: Procedural statistical knowledge test score. **Importance**: After-test importance of procedural statistical knowledge. **Performance**: After-test expectation for performance in procedural statistical knowledge test. Significance level: *, $p < .05$, **, $p < .01$, ***, $p < .001$.

Findings from Table XI and XII showed that participants in the full responsibility group seemed to treat knowledge test scores *independently* from how they rated their beliefs about the importance of conceptual/procedural statistical knowledge; for participants in the knowledge responsibility group the higher of their knowledge test scores, the stronger beliefs about the importance of conceptual/procedural statistical knowledge were reported.

So far, I verified that participants from the 3 responsibility groups reported equivalent baseline commitment to learning statistics and similar academic background characteristics. This step was crucial for making sure the equal variance across groups prior to conducting any responsibility primes in my study. I also found participants in the full responsibility group reported their responsibility beliefs about the importance of conceptual and procedural statistical knowledge *independently* from the results of their conceptual and procedural statistical knowledge score. In contrast, participants in the knowledge responsibility group who scored higher in their conceptual and procedural statistical knowledge test also tended to report stronger beliefs about the importance of conceptual and procedural statistical knowledge.

Testing of Differences in Post-Test Commitment to Learning Statistics at the Midterm

Scores of their commitment to learning statistics were collected by asking participants to complete a post-test survey in the middle of the semester, after their 1st quarter exam and right before their midterm exam. I reported descriptive statistics of their perceptions of persistence decisions, perceived statistical ability, success caused by ability/effort/easy task/luck, and failure caused by low ability/little effort/difficult task/bad luck. As mentioned before, the construct of attitude toward statistics was not included in the final analysis.

Post-Test Attrition

There was about 16% (n = 27) of the participants who didn't complete their post-test surveys. Bias analysis was conducted to examine if there was any significant difference in their baseline scores of the commitment to learning statistics between the stayed-in participants and dropped-out participants (Enders, 2010). A dummy, categorical variable was created for the group of stayed-in and the group of dropped-out. Retrospective analysis was conducted by comparing means of their baseline commitment score in those two groups. Results from the two-tailed t-tests indicated there were no significant differences in their baseline commitment to learning statistics between the stayed-in participants and the dropped-out participants, $ps > .05$ with Bonferroni corrections. Result of Levene's test for equality of variances did not show significant unequal variances.

Among participants who consented to record their statistics exam scores, I did two-tailed t-tests on their exam scores for those who stayed in my study and who dropped out. There was no significant difference in the dropped-out participants and the stayed-in participants' exam scores (See Appendix G Table A1). It indicated there seemed no correlation of participants' statistics exam scores and their completion of the post-test survey. Therefore, I proceeded to the next step of analysis although I was cautious that this attrition of data might be missing not at random.

Post-Test Commitment to Learning Statistics by Responsibility Group

To answer research question 3, MANOVA or MANCOVA was conducted since there were multiple components of individuals' post-test commitment to learning statistics. Three sets of dependent variables were entered in MANOVA/MANCOVA separately: (1) persistence decisions and perceived statistical ability, (2) beliefs about success caused by ability, effort, easy task, and luck, and (3) beliefs about failure caused by low ability, little effort, difficult task, and bad luck. The 1^{st} set represents two common orientations of individuals' commitment; the 2^{nd} and 3^{rd} set represent most common attribution factors for the success and the failure scenario in statistics.

Research question 3. Will undergraduates from the full responsibility, the knowledge responsibility, and the undefined responsibility group report differently on their commitment to learning statistics in the middle of the semester? Will their commitment to learning statistics differ after controlling for previous experience of learning mathematics/statistics?

Answer to research question 3. First, MANOVA was conducted using *persistence decisions* and *perceived statistical ability* as dependent variables and responsibility group as the independent variable. Multivariate test result showed there was no main effect from responsibility groups on persistence decisions and perceived statistical ability. Test of between-subjects effects result indicated that there was a significant difference in persistence decisions among responsibility groups, $F(2, 137) = 3.80$, $p < .05$, $\eta_p^2 = .05$; there was no difference in perceived statistical ability, $F(2, 137) = 1.35$, $p > .05$. The post hoc test for persistence decisions indicated that participants in the full responsibility group reported stronger beliefs ($M = 5.61$) about their persistence decisions than those in the undefined responsibility group ($M = 5.06$), difference = .55, $p < .05$ after Bonferroni corrections. MANOVA was also conducted using *success* attribution factors and *failure* attribution factors and there was no significant difference in any attribution factors among 3 responsibility groups.

Second, considering a covariate model, MANCOVA was conducted using *persistence decisions* and *perceived statistical ability* as dependent variables and the number of high school mathematics/statistics courses as the covariate. Multivariate test result showed there was a main effect from the number of high school mathematics/statistics courses on persistence decisions and perceived statistical ability using Wilks' Lambda, $F(2, 117) = 9.28$, $p < .001$, $\eta_p^2 = .14$. Test of between-subjects effects result indicated that there was no significant difference in persistence decisions and perceived statistical ability among 3 responsibility groups, $F(2, 118) = 2.82$, $F(2, 118) = .76$, $ps > .05$. There was a significant difference in persistence decisions and perceived statistical ability with the number of high school mathematics/statistics courses, $F(1, 118) = 4.51$, $F(1, 118) = 18.68$, $ps < .05$, $\eta_p^2 = .04$ and .14, respectively. Results of MANCOVA indicated that after controlling for the number of high school mathematics/statistics courses, there was no significant difference in persistence decisions and perceived statistical ability among responsibility groups. MANCOVA was also conducted for success attributions and failure attributions and analysis results were the same as using MANOVA.

Testing of Differences in Delayed Post-Test Commitment to Learning Statistics at the Final

At the end of the semester, participants were asked again to report their commitment to learning statistics in another online survey, which was sent to them via their university email. Before answering the research question 4, I reported delayed post-test attrition on individuals' commitment to learning statistics.

Delayed Post-Test Attrition

There was a marked attrition rate in the completion of their final survey at the end of the semester. Among the 140 stayed-in participants who completed their 1st post-test surveys, about 30% (42) of them didn't complete their delayed post-test surveys. The total number of participants who completed their delayed post-test was 98. Bias analysis was conducted (Enders, 2010). A dummy, categorical variable was created for the group of stayed-in and the group of dropped-out in their delayed post-test surveys. Retrospective analysis was conducted by comparing means of their baseline commitment score in those two groups. Results from the two-tailed *t*-tests indicated there were no significant differences in their baseline commitment to learning statistics between the stayed-in participants and the dropped-out participants, *ps* > .05 with Bonferroni corrections. Result of Levene's test for equality of variances did not show significant unequal variances.

Among participants who consented to record their statistics exam scores, I did two-tailed *t*-tests on their exam scores for those who stayed in my study and who dropped out. There was no significant difference in the dropped-out participants and the stayed-in participants' exam scores (See Appendix G Table A2). It indicated there seemed no correlation of participants' statistics exam scores and their completion of the delayed post-test survey. Therefore, I proceeded to the next step of analysis although I was cautious that this attrition of data might be missing not at random.

Delayed Post-Test Commitment to Learning Statistics by Responsibility Group

There was a marked attrition rate in participants' delayed post-test surveys. In spite of the attrition, it is useful to explore whether participants from 3 responsibility groups differ in their report on their commitment to learning statistics at the end of the semester.

Research question 4. Will undergraduates from the full responsibility, the knowledge responsibility, and the undefined responsibility group report differently on their commitment to learn statistics at the end of the semester? Will their commitment to learning statistics differ after controlling for previous experience of learning mathematics/statistics?

Answer to research question 4. First, MANOVA was conducted using *persistence decisions* and *perceived statistical ability* as dependent variables and responsibility group as the independent variable. Result from the multivariate tests showed there was a significant effect from the responsibility group on individuals' persistence decisions and perceived statistical ability using Wilks' Lambda, $F(2, 94) = 4.02$, $p < .05$, $\eta_p^2 = .08$. Test of between-subjects effects demonstrated that there were significant differences in persistence decisions and perceived statistical ability, $F(2, 95) = 4.07$, $F(2, 95) = 4.27$, $p < .05$, $\eta_p^2 = .08$ and .08, respectively. Results of post hoc tests showed (1) participants in the full responsibility group reported stronger persistence decisions than their counterparts in the undefined responsibility group, $p < .05$ with Bonferroni corrections, and (2) participants in the knowledge responsibility group reported stronger perceived statistical ability than their counterparts in the undefined responsibility group, $p < .05$ with Bonferroni corrections.

MANOVA was also conducted using *success attribution* beliefs and *failure attribution* beliefs as dependent variables separately. Result of the multivariate tests indicated no main effect from the responsibility group on reported success and failure attribution beliefs. Result of the test of between-subjects effects showed there was a group difference in success caused by effort, $F(2, 95) = 4.70$, $p < .05$, $\eta_p^2 = .09$. However, Levene's test of equality for error variances indicated there was an unequal error variance of success caused by effort by responsibility groups.

Second, considering a covariate model, MANCOVA was conducted using *persistence decisions* and *perceived statistical ability* as dependent variables and the number of high school mathematics/statistics courses as the covariate. Multivariate test results showed there was a significant effect from the responsibility group on individuals' persistence decisions and perceived statistical ability using Wilks' Lambda, $F(4, 164) = 3.66$, $p < .05$, $\eta_p^2 = .08$. Test of between-subjects effects demonstrated that there were significant differences in persistence decisions and perceived statistical ability, $F(2, 83) = 3.27$, $F(2, 83) = 4.18$, ps < .05, $\eta_p^2 = .07$ and .09, respectively. Results from the post hoc tests showed that (1) participants in the full responsibility group reported stronger persistence decisions than their counterparts in the undefined responsibility group; (2) participants in the knowledge responsibility group reported stronger persistence decisions than their counterparts in the full responsibility and the undefined responsibility group, ps < .05 with Bonferroni corrections.

MANCOVA was also conducted using success attribution beliefs as well as failure attribution beliefs in statistics separately while the number of high school mathematics/statistics courses as the covariate. Again, there was no main effect from the responsibility group on success/failure attribution beliefs. Result of the test of between-subjects effects showed there was a group difference in success caused by effort, $F(2, 83) = 3.39$, $p < .05$, $\eta_p^2 = .08$. Levene's test of equality for error variances indicated an equal error variance. Then, the post hoc test result showed individuals in the full responsibility group reported stronger beliefs about success caused by their effort than their counterparts in the undefined responsibility group, $p < .05$ with Bonferroni corrections.

The findings from delayed post-test confirmed that participants in the full responsibility group tended to report stronger beliefs about their persistence decisions and more likely attribute their success to their effort while their counterparts in the knowledge responsibility group tended to report stronger beliefs about their statistical ability.

Testing of Changes on Commitment to Learning Statistics Across the Semester

The analysis of the within-subjects effects was conducted to test whether participants' perceptions of their commitment to learning statistics would change across the semester. When combining the data from three occasions, it's acknowledged that a smaller sample size ($N = 90$) was obtained when conducting this series of analysis using individuals' commitment to learning statistics, which they reported at the beginning of the semester, at the midterm, and at the final. In each of the following analyses, the focus was on the changes (if any) of each commitment component over time; the responsibility group was entered as the between-subjects factor so that the within-subjects effects could be better explored while partialing out the effect (or "noise") from the between-subjects factor.

Research question 5. Regardless of the responsibility groups, will individual's commitment to learning statistics change over the semester? Will undergraduates from the full responsibility, the knowledge responsibility, and the undefined responsibility group differ in their statistics exam score?

Answer to research question 5. For the first sub-question, a within-subjects design ANOVA was appropriate for exploring the possible changes of participants' commitment to learning statistics across the semester regardless of all 3 responsibility groups. Before detecting any changes, one crucial assumption for conducting such analysis was that repeated-measured variables should be non-independent. Table XIII presents correlation results between persistence decisions, perceived statistical ability, success caused by ability, effort, easy task, and good luck, and failure caused by low ability, little effort, difficult task, and bad luck at the beginning of the semester, at the midterm, and at the final.

Table XIII

Within-Subjects Correlations of Commitment to Learning Statistics Across the Semester

Commitment	Time		
(N = 90)	1 vs. 2	1 vs. 3	2 vs. 3
Persistence	.46***	.38**	.67***
Perceived ability	.61***	.64***	.56***
Success caused by			
Ability	.31**	.44**	.43**
Effort	.40**	.46***	.50***
Easy task	.54***	.44**	.45***
Good luck	.59**	.38**	.43**
Failure caused by			
Low ability	.42**	.40***	.39***
Little effort	.44**	.35***	.31**
Difficult task	.46***	.41**	.44**
Bad luck	.55**	.45***	.41**

***Note*.** **Persistence**: Persistence decisions; **Perceived ability**: Perceived statistical ability. **Time** represented when their commitment to learning statistics was measured. **1**: Time 1 (at the beginning of the semester); **2**: Time 2 (at the midterm); **3**: Time 3 (at the final). Significance level: *, $p < .05$; **, $p < .01$; ***, $p < .001$.

Below I reported the changes of commitment to learning statistics across time in 3 sets of variables: (1) individuals' persistence decisions and perceived statistical ability, (2) attribution beliefs for statistics success scenario, and (3) attribution beliefs for statistics failure scenario.

Persistence decisions and perceived statistical ability. A significant effect on persistence decisions was reported using ANOVA with Greenhouse-Geisser correction, $F(1.80, 158.07) = 3.81$, $p < .05$, $\eta_p^2 = .04$. Marginal means of persistence decisions measured at 3 times are plotted in Figure 3. The test of within-subjects contrasts indicated a significant change: persistence decisions increased at the midterm and decreased at the final (quadratic change). There was no significant effect on perceived statistical ability, $F(2, 174) = 2.10$, $p > .05$. Marginal means of perceived statistical ability measured at 3 times are also plotted in Figure 3.

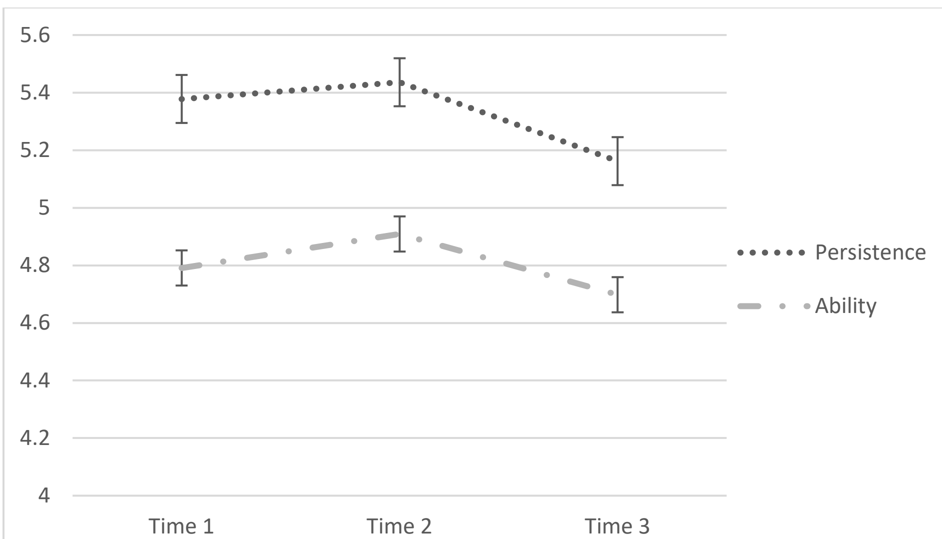

Figure 3. *Estimated marginal means of persistence decisions and perceived statistical ability.*

Attribution beliefs for statistics success scenario. Results from the repeated-measured ANOVA showed there was a significant effect on success caused by ability, $F(2, 174) = 4.85$, $p < .05$, $\eta_p^2 = .05$; There was a significant effect on success caused by effort, $F(2, 174) = 8.85$, $p < .05$, $\eta_p^2 = .09$; There was a significant effect on success caused by easy task, $F(2, 174) = 4.15$, $p < .05$, $\eta_p^2 = .05$. No significant effect was reported on success caused by good luck. Figure 4 shows the trends of attribution beliefs for statistics success across the semester. The test of within-subjects contrasts showed significant changes: success caused by ability decreased at the midterm but increased at the final (indicating quadratic change), while success caused by effort and by easy task decreased all the time (indicating linear change).

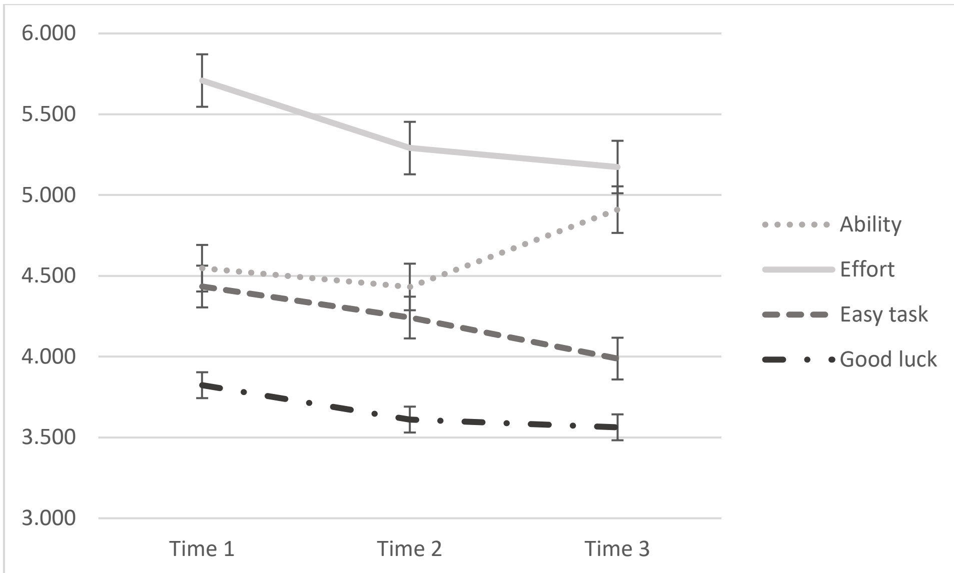

Figure 4. *Estimated marginal means of success caused by ability, effort, easy task, and good luck.*

Among 4 success attribution beliefs, individuals regardless of their responsibility groups tended to report stronger beliefs about success caused by their ability and weaker beliefs about success caused by their effort and by the easiness of statistics task. All belief ratings were above the possible average, which meant that individuals still endorsed (rated positively) all those success attribution beliefs in their statistics course.

Attribution beliefs for statistics failure scenario. Participants' beliefs about statistics failure caused by low ability, little effort, difficult task, and bad luck across the semester were examined using the repeated-measured ANOVA. Individuals' beliefs about *failure caused by low ability*, *little effort, or bad luck* across the semester did not change significantly across the semester, *p*s > .05. However, there was a significant change regarding individuals' beliefs about failure caused by difficult task, $F(2, 174) = 3.52$, $p < .05$, $\eta_p^2 = .04$. The test of within-subjects contrasts showed a significant change: their beliefs about failure caused by difficult task decreased at the midterm but increased at the final (indicating quadratic change).

Figure 5 shows the trend of changes in individuals' attribution beliefs about the failure scenario caused by low ability, little effort, difficulty task, and bad luck. Attribution beliefs about failure caused by low ability, little effort, and bad luck were rated lower than the possible average. However, individuals' beliefs about failure caused by difficult task was always above the possible average. Across the semester, they tended to report stronger beliefs about their statistics failure caused by difficult task in their statistics class.

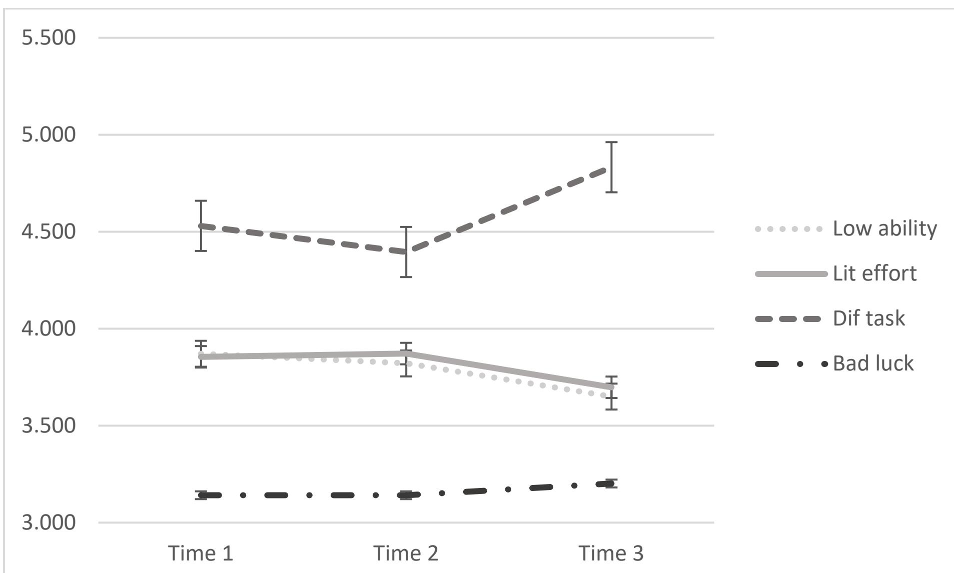

Figure 5. *Estimated marginal means of failure caused by low ability, little effort, difficult task, and bad luck.*

No group difference in statistics exam scores. To answer the second sub-question, participants' raw scores and *T* scores in their 1st quarter, midterm, and final statistics exam are summarized in Table XIV. Since the difficulty of the statistics exams increased across the semester (scores declined in Table XIV), I standardized their raw scores (*z* scores) and transformed them into *T* scores ($T = 50 + 10*z$). ANOVA was conducted using the categorical variable of the 3 responsibility groups and *F* test results showed no significant group difference in their 1st quarter, midterm, and final exam score, *p*s >.05.

Table XIV

Summary of Raw Scores and T Scores in Introductory Statistics Exams

	Full responsibility ($n = 43$)	Knowledge responsibility ($n = 37$)	Undefined responsibility ($n = 38$)
Raw scores			
1st quarter	77.51 (14.80)	77.29 (16.12)	80.05 (12.57)
Midterm	74.05 (16.08)	71.07 (20.00)	76.45 (14.14)
Final	71.00 (18.68)	65.60 (21.80)	70.68 (18.66)
T scores			
1st quarter	49.44 (10.25)	49.28 (11.17)	51.19 (8.71)
Midterm	50.06 (9.58)	48.28 (11.90)	51.49 (8.42)
Final	50.90 (9.50)	48.16 (11.07)	50.76 (9.48)

Predicting Persistence Decisions and Perceived Statistical Ability in the Full Responsibility Group

In this section, I examined the effects of individuals' responsibility for learning statistics on predicting persistence decisions and perceived statistical ability at the midterm and at the final within the full responsibility group. Intention to persist (persistence decisions) and self-confidence in individual's ability (perceived statistical ability) together can offer optimal attributions for engaging students in learning. Regression analysis for the individual differences in persistence decisions and perceived statistical ability by responsibility group was conducted. The focus would be to identify some systematic predictors which could benefit future research.

Research question 6. Among the full responsibility group, will perceived responsibility beliefs from the knowledge primes and the utility value prime predict individuals' commitment to learning statistics in the middle and at the end of the semester?

Answer to research question 6. I reported regression results for persistence decisions and perceived statistical ability at the midterm and at the final. Before that, I reported descriptive statistics of their perceived responsibility beliefs from the knowledge primes and the utility value prime in Table XV. For conceptual statistical knowledge prime and procedural statistical knowledge prime, beliefs about the *importance* of those knowledge primes and *expectation* for performance on those knowledge tests were reported. For the utility value prime, relevance of statistics essays, responsibility for statistics relevance, and responsibility for remembering the value of conceptual and procedural statistical knowledge were reported.

Table XV

Descriptive Statistics of Perceived Responsibility: The Full Responsibility Group

Perceived responsibility	M (SD)
Knowledge prime: Conceptual	
Importance (n = 45)	4.00 (0.98)
Expectation on performance (n = 45)	46.58 (24.34)
Knowledge prime: Procedural	
Importance (n = 40)	3.78 (1.05)
Expectation on performance (n = 40)	55.88 (27.96)
Utility value prime	
Relevance of statistics essays (n = 40)	5.34 (4.2)
Responsibility for statistics relevance (n = 40)	3.65 (0.92)
Remembering the value of conceptual statistical knowledge (n = 38)	4.39 (0.96)
Remembering the value of procedural statistical knowledge (n = 38)	4.48 (1.14)

Persistence decisions at the midterm. Regression was conducted to identify any significant predictors of individuals' persistence decisions at the midterm. Table XVI presents regression statistics for the significant predictors. Notice there was a collinearity issue on responsibility for remembering the utility value of conceptual and procedural statistical knowledge, and therefore, an aggregated score was created and named "responsibility for remembering the utility value of statistical knowledge".

Table XVI

Predictors of Persistence Decisions at the Midterm

Predictors	*Persistence at the midterm*	
	Model 1 Standardized *B*	*Model 2 Standardized B*
Importance of procedural statistical knowledge	.52**	.44**
Remembering the utility value of statistical knowledge		.38**
Adjusted R^2	.25	.38
F	13.31**	11.91***
ΔR^2		.13
ΔF		7.80**

Note. $n = 36$. To simplify the results, regressions with nonsignificant predictors were omitted in this table. Significance level: *, $p < .05$; **, $p < .01$; ***, $p < .001$.

Persistence decisions at the final. Similarly, I followed the previous analytical steps to detect any significant predictors of persistence decisions at the final ($n = 31$). Results showed that their responsibility for perceived utility value of statistics was a significant predictor, standardized $B = .38$; the change of adjusted R^2 was significant, adjusted $\Delta R^2 = .12$, $p < .05$. For participants in the full responsibility group, individuals who reported higher responsibility for perceived utility value of statistics tended to report stronger beliefs about their persistence decisions at the final. As a single predictor, responsibility for perceived utility value explained about 12% of adjusted variance in persistence decisions. Although this finding was not consistent with the finding at the midterm, it could indicate that their responsibility beliefs in the utility value prime turn more evident than their beliefs about the importance of statistical knowledge at the final (since they were not correlated at baseline). It also implied that arguably, both the knowledge primes and the utility value prime contributed to individuals' commitment to learning statistics by raising different types of their responsibility.

Perceived statistical ability at the midterm. Similarly, I followed the previous analytical steps to detect any significant predictors of perceived statistical ability at the midterm ($n = 36$). Regression results showed that only individuals' expectation for performance on procedural statistical knowledge was a significant predictor, standardized $B = .36$, $p < .05$; the change of adjusted R^2 was significant, adjusted $\Delta R^2 = .11$, $p < .05$. Results indicated that for participants in the full responsibility group, individuals who reported stronger beliefs about expectation for performance on procedural statistical knowledge tended to perceive stronger statistical ability at the midterm. This predictor explained about 11% of adjusted variance in perceived statistical ability.

Perceived statistical ability at the final. A similar result was found at the final ($n = 30$): their belief about expectation on performance in procedural statistical knowledge was a significant predictor, standardized $B = .50$, $p < .01$; the change of adjusted R^2 was significant, adjusted $\Delta R^2 = .23$, $p < .05$. Results indicated that participants who reported stronger beliefs about expectation on performance in procedural statistical knowledge tended to perceive stronger statistical ability at the final. This predictor explained about 23% of adjusted variance in perceived statistical ability.

Results in this analysis section confirmed that perceived responsibility for learning among participants in the full responsibility group predicted their persistence decisions as well as their perceived statistical ability later of the semester. In terms of predicting persistence decisions, their belief about importance of statistical knowledge, responsibility for perceived utility value, and responsibility for remembering the utility value of statistical knowledge all played a role. However, in terms of predicting perceived statistical ability, only their belief about expectation on performance in the knowledge tests was significant. Notably, significant predictors in the knowledge primes were all from the procedural statistical knowledge prime. I discussed those findings in the next chapter.

Predicting Persistence Decisions and Perceived Statistical Ability in the Knowledge Responsibility Group

Research question 7. Among the knowledge responsibility group, will perceived responsibility beliefs from the knowledge primes predict individuals' commitment to learning statistics in the middle and at the end of the semester?

Answer to research question 7. Similarly, I reported regression results for persistence decisions and perceived statistical ability at the midterm and at the final. Before that, I reported descriptive statistics of perceived responsibility beliefs in the knowledge primes in Table XVII.

Table XVII

Descriptive Statistics of Perceived Responsibility: The Knowledge Responsibility Group

Perceived responsibility	M (SD)
Knowledge prime: Conceptual	
Importance (n = 42)	3.81 (0.97)
Expectation on performance (n = 42)	46.36 (25.31)
Knowledge prime: Procedural	
Importance (n = 37)	3.94 (0.97)
Expectation on performance (n = 37)	62.92 (29.12)

Persistence decisions at the midterm. Individual's belief about expectation on performance in procedural statistical knowledge was a significant predictor, standardized $B = .46$, $p < .01$. The change of adjusted R^2 was significant, adjusted $\Delta R^2 = .19$, $p < .01$. Results indicated that participants who reported stronger belief about expectation on performance in procedural statistical knowledge test tended to report stronger persistence decisions. This predictor explained about 19% of adjusted variance in perceived statistical ability at the midterm.

Persistence decisions at the final. The similar analysis was conducted using perceived responsibility to predict individuals' belief about persistence decisions at the final (n = 23).

Regression analysis found that only their belief about expectation on performance in procedural statistical knowledge was a significant predictor, standardized $B = .65$, $p < .001$. The change of adjusted R^2 was significant, adjusted $\Delta R^2 = .40$, $p < .01$. Results indicated a consistent significant predictor, and it explained about 40% of adjusted variance in persistence decisions at the final.

Perceived statistical ability at the midterm. Regression analysis was conducted to predict perceived statistical ability at the midterm ($n = 31$). Their belief about expectation on performance in procedural statistical knowledge was a significant predictor, standardized $B = .56$, $p < .05$; the change of adjusted R^2 was significant, adjusted $\Delta R^2 = .29$, $p < .01$. The factor of expectation on performance in procedural statistical knowledge explained about 29% of adjusted variance in perceived statistical ability at the midterm.

Perceived statistical ability at the final. Similarly, the regression analysis was conducted to predict perceived statistical ability at the final ($n = 25$). Here, the importance of conceptual statistical knowledge was a significant predictor, standardized $B = .48$, $p < .01$; the change of adjusted R^2 was significant, adjusted $\Delta R^2 = .20$, $p < .05$. The factor of their belief about the importance of conceptual statistical knowledge explained about 20% of adjusted variance in perceived statistical ability at the final. This was the only finding where expectation on performance was not a significant predictor for the knowledge responsibility group.

In sum, a general pattern seemed to emerge for participants in the full responsibility and the knowledge responsibility group that individuals with stronger expectation on performance in procedural statistical knowledge tended to perceive themselves with higher statistical ability. For participants in the full responsibility group, importance of procedural statistical knowledge, as well as responsibility for perceiving statistics relevance and remembering the utility value of statistical knowledge predicted persistence decisions while for those in the knowledge responsibility group, their expectation on performance in procedural statistical knowledge was still the only significant predictor.

Estimated Retention Rate

Among all participants ($N = 167$), 124 of the respondents consented to record their statistics exam scores along with their research activities. The curriculum policy for the introductory statistics stated that undergraduates who completed their final exam and at least one quarterly exam will be given their final grades. Therefore, the number of participants who completed their final exam and at least one quarterly exam can be an indicator showing that they have retained in their introductory statistics course.

Research question 8. What is the estimated retention rate in the introductory statistics course among all participants?

Answer to research question 8. Among 124 participants who consented to record their statistics exam scores, 118 students had the records of their final exam score and at least one quarterly exam score. The estimated retention rate was 95%. Informed with the introductory statistic course senior lecturer and course coordinator (D Embers, personal communication, January 2020), the introductory statistics course has a general retention rate of 90%. Descriptively, the estimated retention rate in the study sample was slightly higher than the general retention rate in the introductory statistics course.

Chapter summary. Analysis for the baseline commitment scores and academic background information across the 3 responsibility groups showed equivalent variances prior to introducing the responsibility primes**.** Participants in the full responsibility group seemed to treat knowledge test scores *independently* from how they rated their beliefs about the importance of such knowledge tests; this was not the case for participants in the knowledge responsibility group.

I only identified a couple of differences in delayed post-test commitment to learning statistics by responsibility groups at the end of the semester. Participants in the full responsibility group seemed to rate higher on their decisions to persist and more likely attribute their statistics success to their effort. Participants in the knowledge responsibility group seemed to rate higher on their perceived statistical ability.

Across the semester, participants tended to report weaker belief about persistence decisions while maintain the same level of belief about their statistical ability. They tended to report stronger belief about their statistics success caused by their ability, weaker belief about their statistics success caused by their effort, and much stronger belief about their statistics failure caused by the difficult tasks in the statistics class. Their perceived responsibility from the knowledge and utility value primes also predicted their later commitment to learning statistics.

CHAPTER 5

DISCUSSION

Statistics is an important but challenging subject to learn for many undergraduates from various majors. This study discovered that undergraduate students in their introductory statistics course could distinguish conceptual statistical knowledge with procedural statistical knowledge. Both types of knowledge were presumed essential for learning statistics, which is consistent with previous research (Carver et al., 2016; Rittle-Johnson et al., 2015; Star, 2005). Participants in this study believed that they were less familiar with conceptual statistical knowledge (harder for them) and they reported higher perceived importance of such knowledge after completing their knowledge primes. Reportedly, they were more familiar with procedural statistical knowledge and they reported higher expectation for their (future) performance after completing the knowledge primes. The finding in my study justified the application of the K-W-L framework – helping students to realize the difference between what they know and what they *think* they know (Ogle, 1986). Statistics teachers may consider that it would be more beneficial to start with procedural statistical knowledge and then introduce conceptual statistical knowledge (Galli et al., 2011). Evidence from the previous research has supported this approach (Rittle-Johnson et al., 2015).

Participants in the *full responsibility* group demonstrated universally better commitment to learning statistics at the end of the semester. They demonstrated stronger decisions about persistence and strongly believed their success was attributed to their effort in their statistics class. Also, they held positive beliefs about success was attributed to their ability. Individuals in this group demonstrated adaptive attributional beliefs and a growth mindset for learning statistics (Dweck, 2007; Weiner, 1979, 1985). They distinguished the *importance* of conceptual statistical knowledge and procedural statistical knowledge from their actual *performance* in tests of conceptual statistical knowledge and procedural statistical knowledge. This pattern could be helpful for maintaining their responsibility for statistical knowledge while they realize the conceptual statistical knowledge is especially challenging (Carver et al., 2016; Chance, 1997; Cobb, 1992).

Participants in the *knowledge-only responsibility* group reported stronger statistical ability at the end of the semester. Holding a strong belief about statistical ability was not problematic by itself (Bandura, 1997). However, overestimating the importance of the ability could be an issue (Nicholls, 1989). Results from their knowledge primes showed individuals' beliefs about their responsibilities to use conceptual and procedural statistical knowledge were significantly associated with their scores on those tests of such statistical knowledge. Students may feel too preoccupied with feeling good about their ability but overlook the role of their responsibility for learning statistics. Those findings may remind statistics teachers that when students only focused on a knowledge-oriented statistics course they could miss some essential parts of learning statistics.

My study also found that the previous experience in learning mathematics/statistics had a limit effect on undergraduate students' commitment to learning statistics. It was a significant predictor for the commitment to learning statistics in the middle of the semester, but it turned nonsignificant at the end of the semester. Previous research only studied its effect on actual course performance (Hulleman et al., 2010) but not on the individual commitment to learning statistics. Statistics instructors may strategically use statistics learners' previous learning experience to enhance their *continued* commitment to learning the subject matter.

Although I lost a notable proportion of participants at the end of the semester, results from my study demonstrated the benefit of exploring their responsibility for learning statistics. Importance of statistical knowledge and the utility value of statistics, perceived from knowledge and utility value primes, each uniquely predicted individuals' persistence decisions. This indicates a good starting point for exploring a holistic view of undergraduates' responsibility for learning statistics. Arguably, teaching those beliefs and relevant strategies could lead students to alter their beliefs and commitments in academic settings (Yeager & Walton, 2011).

On the other hand, it seems that only their expectation beliefs on their performance in statistical knowledge-related tasks was a predictor of their perceived statistical ability, and those expectation beliefs were predominantly referred to their expectation on the performance of *procedural* statistical knowledge. This could be due to that those participants had more confidence in procedural statistical knowledge, and their expectation and perceived ability (self-efficacy) were highly correlated (Bandura, 1997).

Among the components of participants' commitment to learning statistics, consistent with previous study (Schau & Emmioglu, 2012) individuals tended to report stably perceived statistical ability but weaker persistence decisions across the semester. Their attributional beliefs about success and failure in statistics offered reasonable explanation for their retention. Specifically, stronger perceptions on their beliefs about success caused by their ability and effort and failure caused by difficult tasks enabled individual learners to not blame on themselves about the investment of their ability and effort, and to be realistic and accountable for their course performance. Those findings may inform statistics teachers of notifying students that statistics tasks may be difficult but preservation and adaptive attribution beliefs can help them engage themselves and survive in their statistics class.

One unexpected finding was that participants in the full responsibility group completed the knowledge primes and utility value prime but they did not report stronger statistical ability, as those in the knowledge responsibility-only group. One possible reason could be that the knowledge responsibility questions were too subtle. The questionnaires used for asking responsibility to use statistical knowledge could be revised in future study to directly address their perception of their ability. Task on the relevance of statistics essays may be substituted with directly asking students to reflect and critic their reasons for learning statistics so they could possibly get more engaged (Rosenzweig et al., 2019). Another possible reason could be that participants were fatigued by completing the most complex set of research activities.

Most participants (more than 80%) in this study were freshmen and sophomores. Their learning and retention in the introductory statistics courses may be crucial for setting up their future learning goals and positive expectations for their education in the scientific methodology courses, e.g., statistics courses. Based on my study, almost all participants retained in their introductory statistics course even the course content was perceived difficult by many of them. Their attributional beliefs could function as adaptive strategies for their commitment to learning statistics. Participants in all 3 responsibility groups did not differ in their statistics exam scores. And my study was not designed for implementing any intervention on undergraduate students' statistics performance. When statistics instructors use the information of this study, they should not make causal inference of the association of responsibility and course performance.

My study did demonstrate promising findings on helping students to distinguish conceptual and procedural statistical knowledge as well as to relate statistics to their daily lives (perceiving the utility value of statistics). Participants in the undefined responsibility group also showed adaptive attributional beliefs. It should be mentioned the previous study indicated that the participation in reporting individuals' motivation to learn statistics may affect their retention (Schau & Emmioglu, 2012). Considering the time cost of completing those responsibility primes (about 1 – 1.5 hours), it could be a cost-effective strategy for statistics teachers to remind their students about their responsibility for learning statistical knowledge and adopting flexible attributional beliefs and positive value of statistics.

In the future, I expect to replicate this responsibility framework with a larger sample size and more simplified and explicit measures. Hopefully, this line of research may convey a holistic view that undergraduate students can take their responsibilities from not only distinguishing different types of the domain knowledge but also making meaningful connections between "what students learn" and "why they learn".

REFERENCES

Acee, T. W., & Weinstein, C. E. (2010). Effects of a value-reappraisal intervention on statistics students' motivation and performance. *Journal of Experimental Education, 78*, 487-512. doi: 10.1080/00220970903352753

Ajzen, I., & Madden, T. J. (1986). Prediction of goal-directed behavior: Attitudes, intentions, and perceived behavioral control. *Journal of Experimental Social Psychology, 22*, 453-474. doi: 10.1016/0022-1031(86)90045

Alacaci, C. (2004). Inferential statistics: Understanding expert knowledge and its implications for statistics education. *Journal of Statistics Education, 12*. Retrieved from ww2.amstat.org/publications/jse/v12n2/alacaci.html

Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research, 61*, 315-343. Retrieved from http://www.jstor.org/stable/1170635

Allwood, C. M., & Montgomery, H. (1981). Knowledge and technique in statistical problem solving. *European Journal of Science Education, 3*, 431-450.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369-406. doi: 0.1037/0033-295X.89.4.369

Anderson, J. R. (1985). *Cognitive psychology and its implications* (2nd Ed.). New York: Freeman.

Atkinson, J. W. (1957). Motivational determinants of risk taking behavior. *Psychological Review, 64*, 359-372. doi: 10.1037/h0043445

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

Baroody, A. J., Feil, Y., & Johnson, A. R. (2007). An alternative reconceptualization of procedural and conceptual knowledge. *Journal for Research in Mathematics Education, 38*, 115-131. Retrieved from http://www.jstor.org/stable/30034952

Battle, E. (1965). Motivational determinants of academic task persistence. *Journal of Personality and Social Psychology, 2*, 209-218. doi: 10.1037/h0022442

Bransford, J. D. (1979). *Human cognition*. Belmont, CA: Wadsworth.

Box, G. E. P. (1976). Science and statistics. *Journal of the American Statistical Association, 71*, 791-799. Retrieved from http://www.jstor.org/stable/2286841

Brown, A. L. (1988). Motivation to learn and understand: On taking charge of one's own learning. *Cognition and Instruction, 5*, 311-321. Retrieved from https://www.jstor.org/stable/3233575

Brown, E. N., & Kass, R. E. (2009). What is statistics? *The American Statistician, 63*, 105-110. doi: 10.1198/tast.2009.0019

Bruinsma, M. (2004). Motivation, cognitive processing and achievement in higher education. *Learning and Instruction, 14*, 549-568. doi: 10.1016/j.learninstruc.2004.09.001

Byrnes, J. P., & Wasik, B. A. (1991). Role of conceptual knowledge in mathematical procedural learning. *Developmental Psychology, 27*, 777-786. doi: 10.1037/0012-1649.27.5.777.

Carver, R., Everson, M., Gabrosek, J., Horton, N., Lock, R., Mocko, M., ...Wood, B. (2016). *Guidelines for assessment and instruction in statistics education college report 2016*. Retrieved from http://www.amstat.org/education/gaise

Chance, B. L. (1997). Experiences with authentic assessment techniques in an introductory statistics course. *Journal of Statistics Education, 5*. Retrieved from https://ww2.amstat.org/publications/jse/v5n3/chance.html

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55-81.

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

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1, pp. 7-75). Hillsdale, NJ: Erlbaum.

Chou, C.-P., & Bentler, P. M. (1995). Estimates and tests in structural equation modeling. In R. H. Hoyle (Ed.), Structural equation modeling: Concepts, issues, and applications (pp. 37-55). Thousand Oaks, CA: Sage Publications, Inc.

Chularut, P., & DeBacker, T. K. (2004). The influence of concept mapping on achievement, self-regulation, self-efficacy and in students of English as a second language. *Contemporary Educational Psychology, 29*, 248-263. doi: 10.1016/j.cedpsych.2003.09.001

Cobb, G. (1992). Teaching statistics. In L. A. Steen (Ed.), *Heeding the call for change: Suggestions for curricular action* (pp. 3-43). Washington, DC: Mathematical Association of America.

Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly, 104*, 801-823. doi: 10.2307/2975286

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah, NJ: Erlbaum.

Cohen, S., Smith, G., A., Chechile, R., Burns, G., & Tsai, F. (1996). Identifying impediments to learning probability and statistics from an assessment of instructional software. *Journal Educational and Behavioral Statistics, 21*, 35-54.

Cravalho, P. F. (2010). *Learning statistics using concept maps: Effects on anxiety and performance* (Master's Thesis). Available from ProQuest Dissertations & Theses Global. (Order No. 1482535)

Dauphinee, T. L., Schau, C., & Stevens, J. J. (1997). Survey of attitudes toward statistics: Factor structure and factorial invariance for women and men. *Structural Equation Modeling: A Multidisciplinary Journal, 4*, 129-141. doi: 10.1080/10705519709540066

De Corte, E., & Verschaffel, L. (1987). Children's problem solving skills and processes with respect to elementary arithmetic word problems. In E. De Corte, H. Lodewijks, R. Parmentier, & P. Span (Eds.), *Learning and instruction: European research in an international context* (Vol. I, pp. 297-308). Oxford: Pergamon Press.

de Jong, T., & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist, 31*, 105-113.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

delMas, R. (2004). A comparison of mathematical and statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 79-95). Dordrecht, The Netherlands: Kluwer Academic Publishers.

delMas, R., Garfield, J., Ooms, A., & Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal, 6*, 28-58. Retrieved from http://iase-web.org/documents/SERJ/SERJ6(2)_delMas.pdf

Dochy, F., De Rijdt, C., & Dyck, W. (2002). Cognitive prerequisites and learning: How far have we progressed since Bloom? Implications for educational practice and teaching. *Active Learning in Higher Education, 3*, 265-284. doi: 10.1177/1469787402003003006

Dochy, F., Segers, M., & Buehl, M. M. (1999). The relation between assessment practices and outcomes of studies: The case of research on prior knowledge. *Review of Educational Research, 69*, 145-186. doi: 10.3102/00346543069002145

Dweck, C. S. (2007). *Mindset: The new psychology of success*. New York: Ballantine Books.

Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin, 21*, 215-225.

Eccles (Parsons), J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., … Midgley, C (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75-146). San Francisco, CA: W. H. Freeman.

Emmioglu, E. & Capa-Aydin, Y. (2012). Attitudes and achievement in statistics: A meta-analysis study. *Statistics Education Research Journal, 11*, 95-102. Retrieved from http://iase-web.org/documents/SERJ/SERJ11(2)_Emmioglu.pdf

Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford Press.

Finney, S. J., & Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology, 28*, 161-186. doi: 10.1016/S0361-476X(02)00015-2

Freng, S., Webber, D., Blatter, J., Wing, A., & Scott, W. D. (2011). The role of statistics and research methods in the academic success of psychology majors: Do performance and enrollment timing matter? *Teaching of Psychology, 38*, 83-88. doi: 10.1177/0098628311401591

Galli, S., Chiesi, F., & Primi, C. (2011). Measuring mathematical ability needed for "non-mathematical" majors: The construction of a scale applying IRT and differential item functioning across educational contexts. *Learning and Individual Differences, 21*, 392-402. doi: 10.1016/j.lindif.2011.04.005

Garfield, J. B. (1995). How students learn statistics. *International Statistical Review, 63*, 25-34. doi: 10.2307/1403775

Garfield, J. B. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education, 10*. Retrieved from ww2.amstat.org/publications/jse/v10n3/garfield.html

Garfield, J. B. (2003). Assessing statistical reasoning. *Statistics Education Research Journal, 2*, 22-38. Retrieved from https://iase-web.org/documents/SERJ/SERJ2(1).pdf

Garfield, J. B., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education, 19*, 44-63. doi: 10.2307/749110

Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review, 75*, 372-396.

Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. Dordrecht, The Netherlands: Springer.

Gelman, A. (2005). A course on teaching statistics at the university level. *The American Statistician, 59*, 4-7, doi: 10.1198/000313005X24110

Glaser, R. (1983). *Education and thinking: The role of knowledge* (Report No. PDS-6). Pittsburgh: University of Pittsburgh.

Glaser, R. (1991). The maturing of the relationship between the science of learning and cognition and educational practice. *Learning and Instruction, 1*, 129-144.

Haapala, A, Pietarinen, J., Rautopuro, J., Valtonen, E., & Vaisanen, P. (2002, September). *Concept map as a tool for meaningful learning and assessment in an introductory statistics course*. Paper presented at European Conference on Educational Research, Lisbon, Portugal.

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1-27). Hillsdale, NJ: Lawrence Erlbaum Associates.

Horton, N. J. (2015). Challenges and opportunities for statistics and statistical education: Looking back, looking forward. *The American Statistician, 69*, 138-145. doi: 10.1080/00031305.2015.1032435

Hulleman, C. S., & Barron, K. E. (2016). Motivation interventions in education: Bridging theory, research, and practice. In L. Corno & E. M. Anderman (Eds.), *Handbook of Educational Psychology* (Third Edition, pp. 160-171). Routledge.

Hulleman, C. S., Godes, O., Hendricks, B. L., & Harackiewicz, J. M. (2010). Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology, 102*, 880-895. doi: 10.1037/a0019506

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science, 326*, 1410-1412. doi: 10.1126/science.1177067

Hulleman, C. S., Kosovich, J. J., Barron, K. E., & Daniel, D. B. (2017). Making connections: Replicating and extending the utility value intervention in the classroom. *Journal of Educational Psychology, 109*, 387-404. doi: 10.1037/edu0000146

IBM Corp. (2017). *IBM SPSS Statistics for Windows, Version 25.0*. Armonk, NY: IBM Corp.

Lazowski, R. A., & Hulleman, C. S. (2016). Motivation interventions in education: A meta-analytic review. *Review of Educational Research, 86*, 602-640. doi: 10.3102/0034654315617832

Mathematics education in the United States. (n.d.). Retrived from https://en.wikipedia.org/wiki/Mathematics_education_in_the_United_States

Moore, D. S. (1997). New pedagogy and new content: The case of statistics. *International Statistical Review, 65*, 123-165. doi: 10.2307/1403333

Moore, D. S., & Cobb, G. W. (2000). Statistics and mathematics: Tension and cooperation. *American Mathematical Monthly, 107*, 615-630. Retrieved from http://www.jstor.org/stable/2589117

Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76*, 413-448. doi: 10.3102/00346543076003413

Nicholls, J. G. (1989). *The Competitive ethos and democratic dducation*. Cambridge, MA: Harvard University Press.

Novak, J. D. (1990). Concept mapping: A useful tool for science education. *Journal of Research in Science Teaching, 27*, 937-949.

Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. New York and Cambridge, UK: Cambridge University Press.

Ogle, D. M. (1986). K-W-L: A teaching model that develops active reading of expository text. *The Reading Teacher, 39*, 564-570. Retrieved from http://www.jstor.org/stable/20199156.

Olani, A., Hoekstra, R., Harskamp, E., & van der Werf, G. (2011). Statistical reasoning ability, self-efficacy, and value beliefs in a reform based university statistics course. *Electronic Journal of Research in Educational Psychology, 9*, 49-72. Retrieved from https://eric.ed.gov/?id=EJ926460

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*, 543-578. Retrieved from http://www.jstor.org/stable/1170653

Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*, 429-434. doi: 10.1037/0022-0663.84.4.429

Pollatsek, A., Lima, S., & Well, A. (1981). Concept or computation: Students' understanding of the mean. *Educational Studies in Mathematics, 12*, 191-204. Retrieved from http://www.jstor.org/stable/3482364

Rittle-Johnson, B., & Koedinger, K. R. (2005). Designing knowledge scaffolds to support mathematical problem solving. *Cognition and Instruction, 23*, 313-349. doi: 10.1207/s1532690xci2303_1

Rittle-Johnson, B., Schneider, M., & Star, J. R. (2015). Not a one-way street: Bidirectional relations between procedural and conceptual knowledge of mathematics. *Educational Psychology Review, 27*, 587-597. doi: 10.1007/s10648-015-9302-x.

Roberts, L. (1999). Using concept maps to measure statistical understanding. *International Journal of Mathematical Education in Science and Technology, 30*, 707-717.

Rosenzweig, E. Q., Hulleman, C. S., Barron, K. E., Kosovich, J. J., Priniski, S. J., & Wigfield, A. (2019). Promises and pitfalls of adapting utility value interventions for online math courses. *The Journal of Experimental Education, 87*, 332-352. doi: 10.1080/00220973.2018.1496059

Ruiz-Primo, M. A., & Shavelson, R. J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching, 33*. 569-600. doi: 10.1002/(SICI)1098-2736(199608)33:6<569::AID-TEA1>3.0.CO;2-M

Ruiz-Primo, M. A., Schultz, S. E., Li, M., & Shavelson, R. J. (2001). Comparison of the reliability and validity of scores from two concept-mapping techniques. *Journal of Research in Science Education, 38*, 260-278. doi:10.1002/1098-2736(200102)38:2<260::AID-TEA1005>3.0.CO;2-F

Schau, C. (2003, August). *Students' attitudes: The "other" important outcome in statistics education*. Paper presented at the Joint Statistics Meetings, San Francisco, CA.

Schau, C., & Emmioglu, E. (2012). Do introductory statistics courses in the United States improve students' attitudes? *Statistics Education Research Journal, 11*, 86-94. Retrieved from http://iase-web.org/documents/SERJ/SERJ11(2)_Schau.pdf

Schau, C., & Mattern, N. (1997). Use of map techniques in teaching applied statistics courses. *The American Statistician, 51*, 171-175. doi: 10.1080/00031305.1997.10473955

Schneider, M., & Stern, E. (2010). The developmental relations between conceptual and procedural knowledge: A multimethod approach. *Developmental Psychology, 46*, 178-192. doi: 10.1037/a0016701

Schunk, D. H. (1994). Self-regulation of self-efficacy and attributions in academic settings. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed., pp. 125-151). Mahwah, NJ: Erlbaum.

Schutz, P. A., Drogosz, L. M., White, V. E., & Distefano, C. (1998). Prior knowledge, attitude, and strategy use in an introduction to statistics course. *Learning and Individual Differences, 10*, 291-306.

Schwartz, D. L., Sears, D., & Chang, J. (2007). Reconsidering prior knowledge. In M. Lovett & P. Shah (Eds.), *Thinking with Data* (pp. 319-344). Mahwah, NJ: Erlbaum.

Shepard, L. A., Penuel, W. R., & Pellegrino, J. W. (2018). Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. *Educational Measurement: Issues and Practice, 37*, 21-34.

Skemp, R. R. (1982). Symbolic understanding. *Mathematics Teaching, 99*, 59-61.

Smith, P. T. (1987). Levels of understanding and psychology students' acquisition of statistics. In J. A. Sloboda & D. Rogers (Eds.), *Cognitive processes in mathematics* (pp. 157-168). New York, NY: Oxford University Press.

Sotos, A. E. C., Vanhoof, S., Van den Noortgate, W., & Onghena, P. (2007). Students' misconceptions of statistical inference: A review of the empirical evidence from research on statistics education. *Educational Research Review, 2(2)*, 98-113.

Spatz, C. (2011). *Basic statistics: Tales of distributions*. Boston, MA: Cengage Learning.

Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education, 36*, 404-411. Retrieved from http://www.jstor.org.proxy.cc.uic.edu/stable/30034943

Star, J. R. (2007). Foregrounding procedural knowledge. *Journal for Research in Mathematics Education, 38*, 132-135. Retrieved from http://www.jstor.org.proxy.cc.uic.edu/stable/30034953

Star, J. R., Newton, K., Pollack, C., Kokka, K., Rittle-Johnson, B., & Durkin, K. (2015). Student, teacher, and instructional characteristics related to students' gains in flexibility. *Contemporary Educational Psychology, 41*, 198-208. doi: 0.1016/j.cedpsych.2015.03.001

Tempelaar, D. T., Gijselaers, W. H., & van der Loeff, S. S. (2006). Puzzles in statistical reasoning. *Journal of Statistics Education, 14*, 1-26. Retrieved from http://ww2.amstat.org/publications/jse/v14n1/tempelaar.html

Trautwein, U., Nagengast, B., Nagy, G., Jonkmann, K., Marsh, H. W., & Ludtke, O. (2012). Probing for the multiplicative term in modern expectancy-value theory: A latent interaction modeling study. *Journal of Educational Psychology, 104*, 763-777.

Voss, J. F., Greene, T. R., Post, T. A., & Penner, B. C. (1983). Problem-solving skill in the social sciences. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 165-213). New York: Academic Press.

Weiner, B. (1979). A theory of motivation for some classroom experiences. *Journal of Educational Psychology, 71*, 3-25.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92*, 548-573.doi: 10.1037/0033-295X.92.4.548

Weiner, B. (1993). On sin versus sickness: A theory of perceived responsibility and social motivation. *American Psychologist, 48*, 957-965. doi: 10.1037/0003-066X.48.9.957

Wigfield, A., & Cambria, J. (2010). Expectancy-value theory: Retrospective and prospective. In T. C. Urdan & S. A. Karabenick (Eds.), *The decade ahead: Theoretical perspectives on motivation and achievement* (pp. 35-70). Bingley, UK: Emerald Group Publishing Limited. doi: 10.1108/S0749-7423(2010)000016A005

Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review, 12*, 265-310.

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*, 68-81. doi: 10.1006/ceps.1999.1015

Wigfield, A., Tonks, S. M., & Klauda, S. L. (2016). Expectancy-value theory. In K. R. Wentzel & D. B. Miele (Eds.), *Handbook of motivation at school* (2nd Edition, pp. 55-74). New York: Routledge.

Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review, 67*, 223-265. doi: 10.1111/j.1751-5823.1999.tb00442.x

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81*, 267-301. doi: 10.3102/0034654311405999

Zieffler, A., Garfield, J., Alt, S., Dupuis, D., Holleque, K., & Chang, B. (2008). What does research suggest about the teaching and learning of introductory statistics at the college level? A review of the literature. *Journal of Statistics Education, 16*. Retrieved from http://www.amstat.org/publications/jse/v16n2/zieffler.html

Zimmerman, W.A. & Johnson, G. (2017). Exploring factors related to completion of an online undergraduate-level introductory statistics course. *Online Learning, 21*, 191-205. doi: 10.24059/olj.v21i3.1017

APPENDICES

APPENDIX A

An Explanation of Using Conceptual Knowledge and Procedural Knowledge in Solving Statistical Problem

Example question. An education researcher wants to investigate whether there is a statistical difference between male and female students' mathematics standardized test scores. The data is shown below:

The male group:

ID: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Score: 70, 71, 68, 65, 77, 89, 69, 65, 90, 67

The female group:

ID, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21

Score: 75, 76, 78, 75, 87, 91, 70, 75, 90, 72, 90

Solution (explained using Anderson model, 1982).

Phase 1. Problem representation. A student may read the question and form his/her initial understanding of the given problem. In this phase, the student may encode some key information such as mean difference, two groups, and statistical test. This information gets connected to the individual's statistical knowledge especially their conceptual knowledge on inferential statistics. After recognizing and differentiating (or searching), the student may identify the correct solution: using a two-group t-test for comparing the mean difference.

Phase 2. Knowledge compilation. In this phase, the student needs to apply their procedural knowledge for producing the demanded t-test analysis. A procedure of conducing such inferential test is evoked: step 1, state the null hypotheses and the alternative hypotheses; step 2, confirm that an independent two-group t-test is selected; step 3, in order to calculate the t statistics, the mean and standard deviation for male and female group's mathematics scores need to be calculated; step 4, compare the t statistics with the significance table; step 5, given the significance level (by default, .05), decide whether accept the null hypotheses or reject it (in this case, reject the null hypotheses); step 6, interpret the result: the female group students have higher mathematics scores comparing to the male group students, $p < .05$.

Detailed solution:

Step 1: H0. There is no significant difference between male students' and female students' mathematics test scores.

H1. There is a significant difference between male students' and female students' mathematics test scores.

Step 2: An independent two-group t-test (unpaired) technique is selected. It is assumed two populations have the equal variance in this case.

Step 3: Male group: $N_m = 10$, $M_m = 73.1$, $SD_m = 9.30$;

Female group: $N_f = 11$, $M_f = 79.91$, $SD_f = 7.93$.

$$t = \frac{\overline{Mm} - \overline{Mf}}{s_p \sqrt{\frac{1}{Nm} + \frac{1}{Nf}}}$$

With $$s_p = \sqrt{\frac{(Nm-1)SDm^2 + (Nf-1)SDf^2}{Nm + Nf - 2}}$$

So, we get the $s_p = 8.606$;

Then, we get the t-statistics = -1.81.

Step 4: We take the absolute value of t-statistics and compare it with the t-test significance table (see Figure A1). Given $\alpha = .05$, $df = 19$, the value for a significant t-statistics is 2.093. $1.81 < 2.093$, and therefore, $p > .05$.

Step 5: The comparison result shows it is not statistically significant (given $\alpha = .05$).

Step 6: Interpretation: There is no significant difference between male students and female students' mathematics test scores. In other words, the score variations between two groups are most likely due to sampling error.

DF	A	0.80	0.90	0.95	0.98	0.99	0.995	0.998	0.999
	P	0.20	0.10	0.05	0.02	0.01	0.005	0.002	0.001
1		3.078	6.314	12.706	31.820	63.657	127.321	318.309	636.619
2		1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.599
3		1.638	2.353	3.182	4.541	5.841	7.453	10.215	12.924
4		1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5		1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6		1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7		1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8		1.397	1.860	2.306	2.897	3.355	3.833	4.501	5.041
9		1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10		1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11		1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12		1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13		1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14		1.345	1.761	2.145	2.625	2.977	3.326	3.787	4.140
15		1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073
16		1.337	1.746	2.120	2.584	2.921	3.252	3.686	4.015
17		1.333	1.740	2.110	2.567	2.898	3.222	3.646	3.965
18		1.330	1.734	2.101	2.552	2.878	3.197	3.610	3.922
19		1.328	1.729	2.093	2.539	2.861	3.174	3.579	3.883
20		1.325	1.725	2.086	2.528	2.845	3.153	3.552	3.850
21		1.323	1.721	2.080	2.518	2.831	3.135	3.527	3.819
22		1.321	1.717	2.074	2.508	2.819	3.119	3.505	3.792
23		1.319	1.714	2.069	2.500	2.807	3.104	3.485	3.768
24		1.318	1.711	2.064	2.492	2.797	3.090	3.467	3.745
25		1.316	1.708	2.060	2.485	2.787	3.078	3.450	3.725
26		1.315	1.706	2.056	2.479	2.779	3.067	3.435	3.707
27		1.314	1.703	2.052	2.473	2.771	3.057	3.421	3.690

Figure A1. Value of the t-distribution (two-tailed). Retrieved from https://www.medcalc.org/manual/t-distribution.php

APPENDIX B

1. Consent Form

University of Illinois at Chicago
Research Information and Participation Agreement for
Participation in Social Behavioral Research

Statistics and Motivation

You are being invited to participate in a research study. Researchers are required to provide a consent form such as this one to tell you about the research, to explain that taking part is voluntary, to describe the risks and benefits of participation, and to help you to make an informed decision. You should feel free to ask the researchers any questions you may have.

Principal Investigator Name and Title: Kuan Xing, Doctoral Candidate
Department and Institution: Educational Psychology, University of Illinois at Chicago
Address and Contact Information: 1040 W. Harrison St (MC 147), Chicago IL 60607-7133; phone: (312) 804-3415; E-mail address: kxing2@uic.edu
Why am I being asked?

You are being asked to be a subject in a research study about your views on learning statistics. I am interested in learning more about how undergraduate students think about their statistics knowledge and their motivation to learn it.

You have been asked to participate in this research because you are currently enrolled in an undergraduate-level statistics course.

Your participation in this research is voluntary. A minimum of 150 students and a maximum of 500 students will be involved in this research at UIC. Your decision about whether or not to participate will have no effect on your class standing, your grades, or your current and future relationships at the University of Illinois at Chicago. **If you decide to participate, you are free to withdraw at any time without affecting that relationship.**

What is the purpose of this research?

As you may know, mastering statistical knowledge well enough to think statistically is central to success in many disciplines. As an educational psychologist, I am interested in how statistics learners report their statistical knowledge and motivation to learn statistics. Findings from this research study will be informational, and I hope to strength statistics education.

What procedures are involved?

You will be invited to complete a lab session at UIC or an online session at your convenient time. When you return your consent, please let me know which way you prefer to participate in this study. I will contact you soon by email to confirm with you and let you know the details about finishing your first set of research activities.

You will need to come to the study site once to complete activities in a group setting if you want to do the lab session. When you arrive, you will be randomly assigned into one of the three study conditions. The first condition includes a motivation survey, statistical knowledge questionnaires, statistics relevance task, and a background survey; the second condition includes a motivation survey, statistical knowledge questionnaires, and a background survey; and the third condition includes a motivation survey, a psychological well-being survey, and a background survey. If you want to do the onine questionnaires at your own time, you will receive an online survey. Please read the instructions carefully online. Together, those questionnaires should take about up to one hour to finish. I highly recommend you reach all the questions and do you best to answer all of them.

Right after finishing your first set of research activities, you will be asked to offer your UIC email to receive two short, follow-up motivation surveys, and each of which should take about five minutes or so to finish. The first one may be sent to you by email in about 2-3 weeks after you have done the initial research activities. The last one may be sent to you by email in April. At the end of the semester, you will be given the access to all the research materials included in this study. The research materials will include statistical knowledge questionnaires, motivation surveys, statistics relevance task, the background survey, and the psychological well-being survey. You will have the access to all study materials and will receive a debriefing email which explains the purpose of the study.

You can choose not to give the permission to link to your statistics exam scores and still participate in the study. In this case, your data will be included in the final analysis and group-level statistical analyses will be conducted. You will be asked at the end of this consent if you still want your data to be collected and used. You have the rights to give the access or not.

Finally, after your instructor has entered your final grade for this course, I will use your consent form to collect your exam scores, and to link them to your study responses. After completing this step, I will permanently remove all of the information that identifies you, including your name and your UIC email address.

What are the potential risks and discomforts?

To the best of my knowledge, the things you will be doing have no more risk of harm than you would experience in everyday life. A risk of this research is a loss of privacy (revealing to others that you are taking part in this study) or confidentiality (revealing information about you to others to whom you have not given permission to see this information). During the study recruitment, your classmates may know that you participate in this study. Statistics lecturers will only know you participated when I approach them to collect study volunteers' statistics exam scores after your final grade for the course has been submitted.

Once in a while, someone finds some of the research tasks uncomfortable. If that happens to you, you can choose not to answer questions or decide to stop participating in this research and there will be no penalty for your decision. I need complete answers to all our questions for the study, but can simply take your ideas out of the project.

As part of the procedure, you will be asked to give me access to your statistics exam scores and your UIC email address. Some of you may have concerns about privacy. I would like to ensure you that I will remove all personal information once your final grade for the course has been submitted and linked to your other data. Findings will be reported based on aggregated analysis of de-identified information. Of course, you have the right to not allow access to this information and there is no penalty at all. But the information you offer will really help me understand better about my research questions.

Statistics lecturers will only know who has agreed to participate in the study when I ask for statistics exam scores after final grades have been submitted. This aspect of the research data will be saved on an encrypted laptop and locked in my advisor's office, but all identifiers will be deleted after the statistics exam scores are matched with your other study responses.

Finally, the surveys' answers in the computer session will be collected using software called Qualtrics. This software also reports answers in an encrypted format. Your answers will be tracked using an ID number rather than any personal information about you. This ID allows me to connect your computer session answers to your paper session answers. In my analysis, I will combine those answers and analyze them at the group level. I will do everything I can to protect your privacy on my end, but because Qualtrics is not owned by me, I cannot guarantee that they may not use your answers for a purpose I do not endorse – fortunately, they have promised that they will honor the security agreements UIC has established with them and not share the research data with anyone else.

Are there benefits to taking part in the research?

There is no direct benefit to you for participating this research study. However, by collecting your responses and feedback, educational psychologists may learn new things about helping other students. In particular, I would like to understand better how statistics learners view their motivation in learning statistics.

What about privacy and confidentiality?

Research staff will know you participated in this research and, after your final grade for this course has been submitted, so will your instructor. Other students in the course and/or in the lab session will know who has participated. None of your responses to the research activities will be disclosed to others without your written permission, unless such disclosure is necessary to protect your rights or welfare (for example, if the UIC Office for the Protection of Research Subjects monitors the research or consent process) or if required by law.

Copies of your signed consent form and the identifying information (UIC email address) will be temporarily stored on an encrypted portable hard drive that will be locked in my advisor's office. All information will be encrypted using the DiskCryptor software. Your name, email, and any identifying information will be permanently destroyed following data collection, including the collection of your exam scores after your final grade for this course has been submitted. Paper copies of the UIC email address that you provide at the end of this consent form will be pulled off and shredded as soon as possible following data collection.

Only research staff, representatives from the Office for the Protection of Research Subjects at UIC, and I will have access to all aspects of this project. When the results of the research are published or discussed in conferences, no information will be included that would reveal your identity.

What are the costs for participating in this research?

There are no costs to you for participating in this research.

Can I withdraw or be removed from the study?

If you decide to participate, you are free to withdraw your consent and discontinue participation at any time. You just need to tell me that you want to stop. I also have the right to stop your participation in this study without your consent. I would do that if I believe it is in your best interest because I notice that you are experiencing an unusual amount of stress as you answer the questions. If you signed the consent form but then decided to withdraw, you will still be included in the gift card lottery.

Who should I contact if I have questions?

Contact the principal investigator Kuan Xing via email kxing2@uic.edu or by phone (312)-804- 3415 or the faculty sponsor Dr. Theresa (Terri) Thorkildsen, professor, via email thork@uic.edu or by phone (312)-996-8138 if you have any questions.

What are my rights as a research subject?

If you feel you have not been treated according to the descriptions in this form, or if you have any questions about your rights as a research subject, including questions, concerns, complaints, or to offer input, you may call the Office for the Protection of Research Subjects (OPRS) at 312- 996-1711 or 1-866-789-6215 (toll-free) or e-mail OPRS at uicirb@uic.edu.

What if I am a UIC student?

You may choose not to participate or to stop your participation in this research at any time. This will not affect your class standing or grades at UIC or your current and future relationship with lecturers and UIC. The investigator may also end your participation in the research. If this happens, your class standing or grades will not be affected. You will not be offered or receive any special consideration if you participate in this research.

Remember:

Your participation in this research is voluntary. Your decision about whether or not to participate will not affect your current or future relations with the University. If you decide to participate, you are free to withdraw at any time without affecting that relationship.

<u>Signature of Subject</u>

Do you agree to participate in this study?

☐ Yes

☐ No

Do you agree to give me your UIC email address and let me access your statistics exam score?

☐ Yes, I would like to offer the above information and give the access to my statistics exam score.

☐ No, I would not agree to give the access to my statistics exam score; but Yes I do give permission for my data to be collected and used for checking measurement tools' reliability and validity and for conducting group-level statistical analyses.

I have read (or someone has read to me) the above information. I have been given an opportunity to ask questions and my questions have been answered to my satisfaction. I agree to participate in this research. I will be given a copy of this signed and dated form if needed.

________________________________ _________________

Signature Date

____

Printed Name

________________________________ _________________

Signature of Person Obtaining Participation Agreement Date (must be same as subject's)

____

Printed Name of Person Obtaining Participation Agreement

If you said yes to the above question, please offer your UIC email address and UIN below (print):

_________________@uic.edu

UIN: ___________________

2. IRB Approval Letter

Approval Notice

Initial Review (Response To Modifications)

April 11, 2018

Kuan Xing, BS,MS

Educational Psychology

3014 S Poplar Ave Apt2

Chicago, IL 60608

Phone: (312) 804-3415 / Fax: (312) 996-5651

RE: Protocol # 2017-1199

"Improving Undergraduates' Prior Knowledge and Motivation to Enhance Statistics Performance"

Dear Mr. Xing:

Please note that stamped .pdfs of all approved recruitment and consent documents have been uploaded to OPRSLive, and can be accessed under "Approved Documents" tab. Please remember to use only those approved documents to recruit and enroll subjects into this research project. OPRS/IRB no longer issues paper letters or stamped/approved documents.

Your Initial Review (Response To Modifications) was reviewed and approved by the Expedited review process on April 11, 2018. You may now begin your research

Please note the following information about your approved research protocol:

Protocol Approval Period: April 11, 2018 - April 11, 2019

Approved Subject Enrollment #: 500

Additional Determinations for Research Involving Minors: The Board determined that this research satisfies 45CFR46.404, research not involving greater than minimal risk. Therefore, in accordance with

45CFR46.408, the IRB determined that only one parent's/legal guardian's permission/signature is needed. Wards of the State may not be enrolled unless the IRB grants specific approval and assures inclusion of additional protections in the research required under 45CFR46.409. If you wish to enroll Wards of the State contact OPRS and refer to the tip sheet.

Performance Sites: UIC

Research Protocol(s):

a) IRA: Improving Undergraduates' Prior Knowledge and Motivation to Enhance Statistics Performance, Version 5, 04/09/2018

Recruitment Material(s):

a) Group Undergraduate Statistics Class Recruitment Script, Version 2, 01/03/2018
b) The Study Session Schedule Survey, no footer

Informed Consent(s):

a) Consent, Version 5, 04/09/2018

Parental Permission(s):

a) A waiver of parental permission has been granted under 45 CFR 46.116(d) and 45 CFR 46.408(c); for minor college students in the statistics class (minimal risk; 16-17 year old college students in statistics class only; participation is voluntary and otherwise confidential and obtaining parental permission would present intrusion and potential risk of a breach of subject privacy).

Your research meets the criteria for expedited review as defined in 45 CFR 46.110(b)(1) under the following specific category(ies):

(5) Research involving materials (data, documents, records, or specimens) that have been collected, or will be collected solely for nonresearch purposes (such as medical treatment or diagnosis)., (7) Research on individual or group characteristics or behavior (including but not limited to research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies.

Please note the Review History of this submission:

Receipt Date	Submission Type	Review Process	Review Date	Review Action
11/02/2017	Initial Review	Expedited	11/13/2017	Modifications Required

01/22/2018	Response To Modifications	Expedited	02/07/2018	Modifications Required
02/20/2018	Response To Modifications	Expedited	03/07/2018	Modifications Required
03/23/2018	Response To Modifications	Expedited	04/11/2018	Approved

Please remember to:

→ Use your **research protocol number** (2017-1199) on any documents or correspondence with the IRB concerning your research protocol.

→ Review and comply with all requirements on the guidance,

"UIC Investigator Responsibilities, Protection of Human Research Subjects" (http://research.uic.edu/irb/investigators-research-staff/investigator-responsibilities).

Please note that the UIC IRB has the prerogative and authority to ask further questions, seek additional information, require further modifications, or monitor the conduct of your research and the consent process.

Please be aware that if the scope of work in the grant/project changes, the protocol must be amended and approved by the UIC IRB before the initiation of the change.

We wish you the best as you conduct your research. If you have any questions or need further help, please contact OPRS at (312) 996-1711 or me at (312) 413-1518. Please send any correspondence about this protocol to OPRS at 203 AOB, M/C 672.

Sincerely,

Alma Milat, BS

IRB Coordinator, IRB # 2

Office for the Protection of Research Subjects

Enclosure(s): Following approved recruitment and consent documents have been uploaded under "approved documents" tab in OPRSLive:

1. **Informed Consent Document(s):**
    a) Consent, Version 5, 04/09/2018
2. **Recruiting Material(s):**
    a) Group Undergraduate Statistics Class Recruitment Script, Version 2, 01/03/2018
    b) The Study Session Schedule Survey, no footer

cc: Stacey S. Horn, Educational Psychology, M/C 147
Theresa Thorkildsen, Educational Psychology, M/C 147

APPENDIX C

Commitment to Learning Statistics

Q1. How do you think of your commitment to learning statistics? You may choose one number among a 7-point scale where **(1)** represents "**Do not agree at all**", **(4)** represents "**Neutrally**" (neither agree nor disagree), and **(7)** represents "**Totally agree**". **There is no right or wrong answer**.

Please reflect on your **current experience** in your statistics class/learning statistics:	Not agree at all			Neutrally			Totally agree
I intend to learn more in my statistics course.	1	2	3	4	5	6	7
I try to find out the meaning of statistics.	1	2	3	4	5	6	7
I make lots of efforts in learning statistics even sometimes it's challenging.	1	2	3	4	5	6	7
When I study statistics, I put my best foot forward.	1	2	3	4	5	6	7

Note. Exploratory factor analysis for persistence decisions (Q1): One factor (Eigenvalue = 2.60) was comprised of 4 items and explained 65% of variance with factor loading from 0.64 to 0.69. $N = 167$.

Q2. My ability in statistics class is … (Select one)

- Excellent
- Very good
- Moderately good
- Average
- Moderately poor
- Poor
- Very poor

Q3. Comparing to most other students, my ability in statistics class is … (Select one)

- Top of the class
- Top 10% of the class
- Above the class average
- Middle of the class
- Below the class average
- Bottom 10% of the class

- Bottom of the class

Q4. My expectation on the statistics course achievement is … (Select one)

- Excellent
- Good
- Above average
- Average
- Below average
- Poor
- Very poor

Note. Exploratory factor analysis for perceived statistics ability (Q2-Q4): One factor (Eigenvalue = 2.35) was comprised of 3 items and explained 78% of variance with factor loading from 0.68 to 0.84. *N* = 167.

Q5. Suppose you experience the following situations in your statistics course. Among possible reasons, to what extent will you agree with each following statement? If you have your own reasons, please specify and rate below.

Part 1. Assume you perform well in your statistics class. The possible reasons may be:	Not agree at all			Neutrally			Totally agree
My statistics ability is outstanding.	1	2	3	4	5	6	7
I make a lot of efforts to study statistics.	1	2	3	4	5	6	7
The tasks are easy.	1	2	3	4	5	6	7
Good luck.	1	2	3	4	5	6	7
Other (please specify):	1	2	3	4	5	6	7

Part 2. Assume you perform poorly in your statistics class. The possible reasons may be:	Not agree at all			Neutrally			Totally agree
My statistics ability is below average.	1	2	3	4	5	6	7
I make little effort to study statistics.	1	2	3	4	5	6	7
The tasks are difficult.	1	2	3	4	5	6	7

Bad luck.	1	2	3	4	5	6	7
Other (please specify):	1	2	3	4	5	6	7

Q6. On a scale of 0 - 100, how do you like statistics? Write down an integral number (e.g., 60, 75, etc.) which best represents your choice: ___________________

Note. In the 1st post-test survey, Q1's question stem will be written as "You have studied statistics for a while. What do you think of your commitment to learning statistics now?". In the delayed post-test survey, Q1's question stem will be written as "You have studied statistics for a semester. What do you think of your commitment to learning statistics now?". The question stems for the rest of the scale was adjusted accordingly in the 1st post-test and the delayed post-test surveys.

APPENDIX D

1. Rudimentary Conceptual Statistical Knowledge Questionnaire

Part 1. Perceived responsibility for conceptual statistical knowledge

1.1 Before-test questions

Q1. Approximately, what percentage of correct responses you will expect to have in the task #1? Circle one point from the scale of 0% - 100%.

|_____|_____|_____|_____|_____|_____|_____|_____|_____|_____|

0% 10 20 30 40 50 60 70 80 90 100%

Q2. Do you think the concept map about "probability" and related concepts important to learn? (Check one)

- o Extremely important
- o Very important
- o Moderate important
- o Slightly important
- o Not important at all

1.2 After-test questions

<u>Share Your Reflection</u>.

Q3. What is the estimated percentage of your correct responses in the task #1? Circle one point from the scale of 0% - 100%.

|_____|_____|_____|_____|_____|_____|_____|_____|_____|_____|

0% 10 20 30 40 50 60 70 80 90 100%

Q4. Do you think the conceptual knowledge about probability (and related concepts) important to learn? Check one.

- o Extremely important
- o Very important
- o Moderate important
- o Slightly important
- o Not important at all

Part 2. Concept map

Concept mapping is a common-used technique for assessing learners' conceptual knowledge. It is a **virtual way** to **organize/summarize** the conceptual knowledge around the topic(s). Whether you know or don't know about the concept mapping, that's totally fine. You will see a specific example soon.

Example. For instance, you are learning the concept "**food**". It can have many features. Suppose you want to focus on food "**nutrients**". Several concepts are included in the category of food nutrients. To organize your knowledge, you may construct a concept map about "food" (focusing on nutrients & related concepts).

There are several steps to complete a concept map:

(1) **Concepts** are always in your concept map, in this case, *food*, *nutrients*, and categories of food nutrients such as *fiber*, *proteins*, *carbohydrates*, *fats*, *vitamins & minerals*, and *water*.

(2) **Interrelations** between concepts. In other words, in your concept map, you will use specific language to describe/define/discover the interrelations between two concepts. For example, food *contains* nutrition that is good for people's physical health. And one type of nutrition, e.g., proteins, *is good for* cell growth. Those interrelations are called ***links***.

(3) **Hierarchical structure**. Usually, the concept map is organized hierarchically: upper level represents more abstract concepts while lower level represents more specific concepts.

(4) **Examples** (*optional*). This step is optional. In some concept maps, examples may appear at the bottom to explain a concept. For instance, *vegetables & fruits* are examples of the type of food which contains vitamins & minerals.

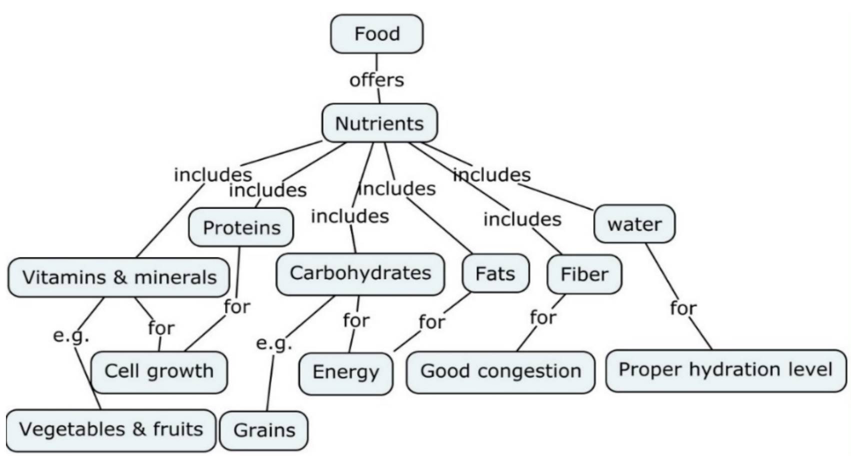

Figure 1. An example of concept map on "food"

Task #1.

Look at Figure 2. They are statistics concepts/statements and links which connect different concepts/statements. The central concept is focusing on "**probability**". Your task is to **complete the blanks** (e.g., N1-7, L1-7) in the incomplete concept map. **Select appropriate options** from the following table. For N1-N7, you select options from A-G. For L1-L7, you select options from H-N. Each option from the list should be used **once** in one blank. You are encouraged to **use your best guess**. You are encouraged to complete all questions.

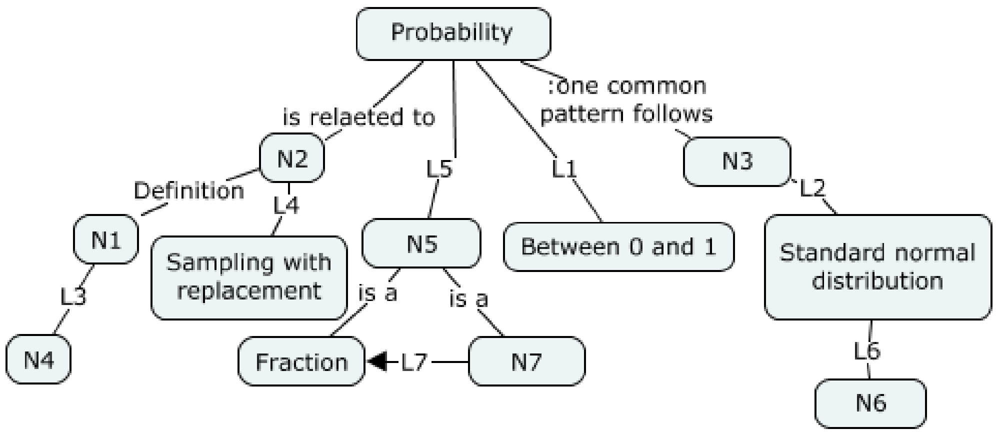

Figure 2. The concept map of probability.

Here are the answer options:

N1-N7 Options:	**L1-L7 Options:**
A. Comparing scores (z-scores) from different normal distributions	**H**. can be described by
B. Each individual in the population has an equal chance of being selected	**I**. fulfill the condition of
C. No bias	**J**. insuring
D. Proportion	**K**. can be observed as
E. Random Sampling	**L**. ranges
F. The Normal Distribution	**M**. can be converted to
G. The observed probability of any particular outcome that can happen when several different outcomes are possible	**N**. is for

YOUR TASK: Fill in N1-N7 and L1-L7 with most appropriate answers (each option used once).

Answer sheet:

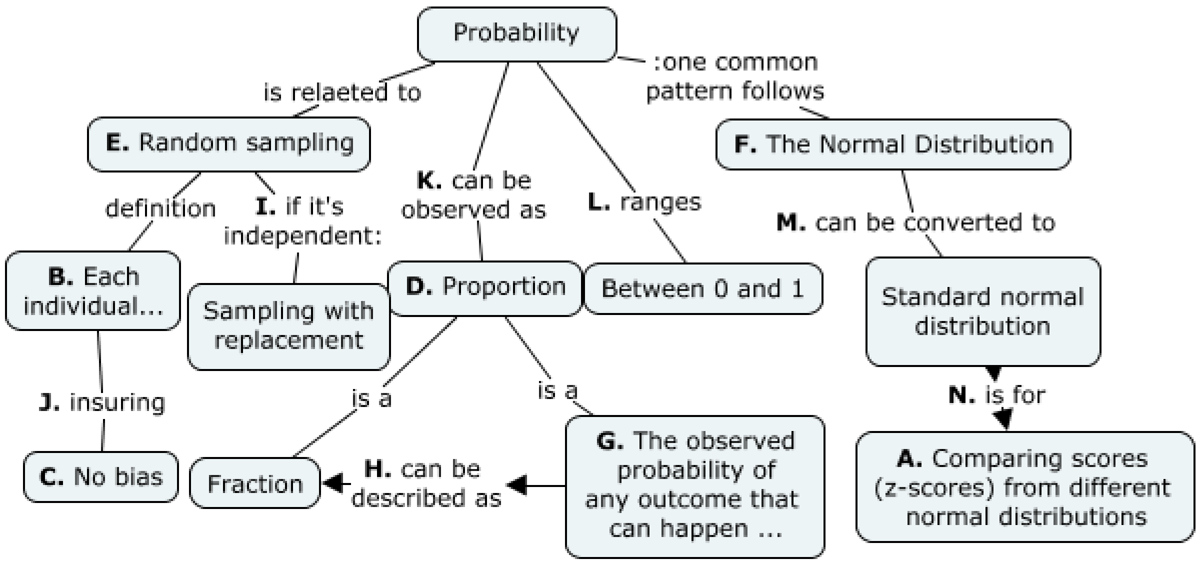

Figure 3. Standard answers to the questions in task #1

FEEDBACK

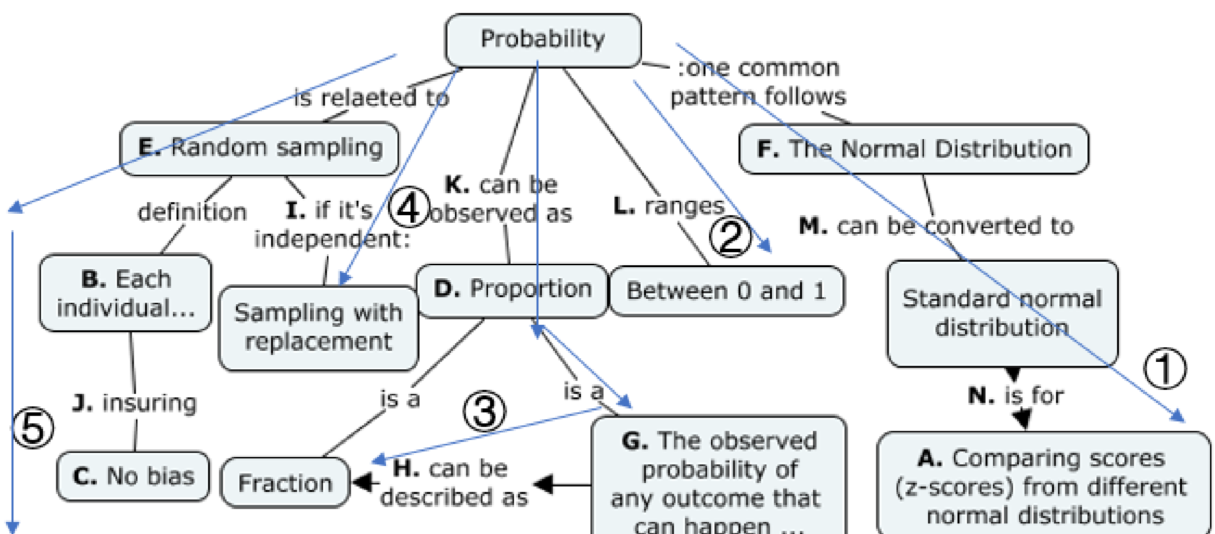

Figure 4. A flowchart for explaining statistical concepts and their interrelations

Please **read carefully the explanations** below. There are five hierarchical networks (P stands for "Probability"): ①P – F – M – N – A, ②P – L, ③P – K – D – G – H, ④P – E – I, and ⑤P – E –

B – J – C. Each time, one hierarchical network will be explained. The explanation starts from the right-hand side:

① P – F – M – N – A:

P. probability. Probability describes how likely an event may happen. In mathematical language, it is the number of ways an event can happen divided by the total number of all possible outcomes.

F. The normal distribution. Distribution is a term to describe the spread of a given data. The normal distribution is one most common distributions you will see in the research as well as in daily life. People's heights, students' SAT/ACT scores, and blood pressure are some examples that follow the normal distributions.

M. Can be converted to (standard normal distribution). See the explanation below:

A. Comparing scores (z-scores) from different normal distributions. First, a z-score is called a standardized score. Accordingly, it follows a standard normal distribution (a normal distribution without a mean = 0 and a standard deviation = 1). Scores converted as standardized scores can be compared with each other since they're on the same scale (same center & unit).

N. is for. Let's check the whole sentence "Standard normal distribution is for comparing scores (z-scores) from different normal distribution". Does it make sense now? Hope so!

② P – L:

L. ranges. It is not difficult to understand that the range of the probability is between 0 and 1. Zero means an event will never happen. One means an event will always happen. In most cases, the probability will be larger than 0 and less than 1.

③ P – K – D – G – H:

K. can be observed as & *D. proportions*. A proportion is a percentage or fraction that is calculated as follows: the count number of a certain attribute divided by the total number of all possible attributes that may present. In general, a probability is a ***theoretical*** term whereas a proportion is an ***empirical*** (observed) term.

H. the observed probability of any particular outcome that can happen when several different outcomes are possible. As you may realize, H is another way to describe the term proportion: based on a finite number of observations, the observed probability of a specific outcome is defined as a proportion among all possible outcomes.

G. can be described as. Fractions and proportions are different formats of quantitative representations about the world. They can be viewed as exchangeable terms in statistics to describe uncertainty.

④ P – E – I:

E. random sampling. A *population* includes all the elements from a set of events. A *sample* includes one or more elements drawn from the population. Random sampling is a means of drawing samples from the given population. Random sampling prevents potential bias when people want to draw inferences about the population from the sample.

I. if it's independent: (sampling with placement). "If it's independent" means that an event happens many times, and one-time result will not affect another-time result. This represents a random sampling.

⑤ P – E – B – J – C:

B. Each individual in the population has an equal chance of being selected & *J. insuring* & *C no bias*. Random sampling is characterized as each individual in the population has an equal chance to be selected (aka, no bias). Why "no bias" matters? If a large number of unbiased random samples is drawn from the population, the average sample could accurately represent the characteristic of the population.

2. Rudimentary Procedural Statistical Knowledge Questionnaire

Part 1. Perceived responsibility for procedural statistical knowledge

1.1 Before-test questions

Q1. Approximately, what percentage of the questions you will expect to solve correctly? Check one point from the scale of 0% - 100%.

|_____|_____|_____|_____|_____|_____|_____|_____|_____|_____|

0% 10 20 30 40 50 60 70 80 90 100%

Q2. Do you think the procedural knowledge important to learn? Check one

- o Extremely important
- o Very important
- o Moderate important
- o Slightly important
- o Not important at all

1.2 After-test questions

<u>Share Your Reflection.</u>

Q3. Imagine those previously similar questions are assigned to you in your later class, what percentage of correct responses you will expect to have? Circle one from the scale of 0% - 100%.

|_____|_____|_____|_____|_____|_____|_____|_____|_____|_____|

0% 10 20 30 40 50 60 70 80 90 100%

Q4. Do you think the procedural knowledge important to learn in the future? (Check one)

- o Extremely important
- o Very important
- o Moderate important
- o Slightly important
- o Not important at all

Part 2.

Problem #1. Solve the following expression: $(17\text{-}13)^2 + (10\text{-}13)^2 + (13\text{-}13)^2 + (12\text{-}13)^2$

Write down your steps for solving this problem below.

Problem #2. Calculate the probability: Person A and person B are practicing archery (which uses a bow to shoot arrows) independently. Based on their records, A has a probability of 0.6 to score 7 or above. B has a probability of 0.5 to score 7 or above. On their next shoot, what is the expected probability that A and B both score 7 or above?

Write down your steps for solving this problem below.

Hint. Denote P (A) = the probability that A scores 7 or above, P (B) = the probability that B scores 7 or above, and P (AB) = the expected probability that A and B both score 7 or above.

P (AB) = ?

Problem #3. Complete: If x/y = 0.25, then y/x= ?

Write down your steps for solving this problem below.

Problem #4. A researcher wants to study the frequency of cellphone use among undergraduates. Ten undergraduates' frequencies of checking them cellphones during one random hour at a random day were recorded. Data is listed below. The researcher wants to know the average frequency of cellphone use in this group of students. If i represents the ith student, x_i represents the ith student's frequency, and n represents the number of students. Please use the following expression to calculate:

The average frequency $= \frac{\sum x_i}{n} = \frac{\sum(\qquad)}{n} = ?$

ID	01	02	03	04	05	06	07	08	09	10
Freq.	1	6	3	0	2	3	4	6	5	3

Write down your steps. Keep two decimals for your final answer.

Problem #5. A teacher has conducted two quarter exams and she found that 26% of the class passed both exams and 40% of the class passed the 2nd quarter exam. Now you are supposed to calculate a "conditional probability": Given the students who passed 2nd quarter exam, what percent of students who also passed the 1st quarter exam?

Hint. Conditional probability is the probability of one thing being true **given** that another thing is true. Suppose there are two events A and B, and they are dependent. You want to know: given B happens, what is the probability of that A also happens? The formula to calculate the conditional probability is $P(A \mid B) = P(A \cap B) / P(B)$. $P(A \cap B)$ represents the probability of that event A and B happen at the same time; P(B) represents the probability of that event B happens. Use this formula to solve the problem.

Write down your steps. Keep two decimals for your final answer. For your convenience, specify event A: students who passed the 1st quarter exam; event B: students who passed the 2nd quarter exam; P (A ∩ B): the probability that students passed both quarter exams;

The question is to calculate: P (A | B) = P (A ∩ B) / P(B) = (please continue)

Answer sheet:

Answer to Problem #1.

Final answer: 26 Your answer: ___ Correct: Yes ___ No ____

Steps	Rules & explanations
The above equation = $4^2 + (-3)^2 + 0^2 + (-1)^2$	Following the order of operations, the numbers in the parentheses have the priority to be calculated.
$4^2 + (-3)^2 + 0^2 + (-1)^2 = 16 + 9 + 0 + 1$	Then, you want to calculate the squares by following the order of operations.
16 + 9 + 0 + 1 = 26	Finally, you just do a simple addition four times and get the final answer.

Note. The order of operations is crucial when you solve equations or do calculations. Here is a simple way to understand different situations: Below (Table 1) is a table of priority levels regarding order of operations. Basically, you want to follow the left-to-right calculation order. However, if any level of operation in the following table is involved, the higher-level operation has the priority to be computed than the lower-level operation.

Table 1. The priority levels of order of operations

Level 1	Level 2	Level 3	Level 4
+, -	*, /	x^n, $\sqrt[n]{x}$	Parentheses

Two simple examples are given below. Calculate the following two examples: (You don't need to really calculate them. You are encouraged to read through the worked examples and self-check if you follow the same order of operations)

Example 1. Complete: $1 - 2 * 3 + 4^2 =$

Answer. Level 1, 2, and 3 was involved. Step 1: calculate exponent term; Step 2: calculate the multiplication term; Step 3: calculate the addition/subtraction terms following the left-to-right order (same level). $1 - 2 * 3 + 4^2$ =1 – 2 *3 + 16 = 1 – 6 + 16 = -5 + 16 = -11

Example 2. Complete: $1 - (2 * 3) * 3^2 =$

Answer. Level 1, 2, 3, and 4 was involved. Step 1: calculate the value in the parentheses; step 2: calculate the exponent value; step 3: calculate the product term; step 4: complete the addition and subtraction calculations following the left-to-right order. $1 - (2 * 3) * 3^2$ = 1 – 6 * 3^2 = 1 – 6 * 9 = 1 – 54 = -53

Answer to Problem #2.

Final answer: 0.3 Your answer: ____ Correct: Yes ___ No ____

Explanation. A's archery scores are independent of B's archery scores. There is a multiplication rule in probability theory: If there are two independent events, P and Q, and their probabilities of occurrence are p and q, the probability of that P and Q both happen is: p * q.

Steps	**Rules & explanations**
Given the information in the question stem, P (A) = 0.6, P (B) = 0.5, P (AB) = P(A) * P(B) [multiplication rule]	Using the multiplication rule above, our first step is to identify P (A), P (B), and P (AB).
P (AB) = P(A) * P(B) = 0.6 * 0.5 = 0.3	Finish the multiplication. Note. If you got 0.6 + 0.5 = 1.1, you can self-check if your answer is correct. Recall that previously we mention: probability is between 0 ~ 1. So, if you have a number larger than 1, it should not be a correct probability. Make sense?

Answer to Problem #3.

Final answer: 4 Your answer: ____ Correct: Yes ___ No ____

Explanation. On the left-hand side, the step-by-step solution will be shown. On the right-hand side, the rules and the explanations will be shown.

Steps	**Rules & explanations**
x/y = 1/4 --> (x/y)*y = (1/4)*y --> x = y/4	Our ultimate goal is to get y/x. So our first step is to multiple y on both sides of the equation. This follows the multiple/division rule for equations. As a reminder, the rule is shown below.
x = y/4 --> x*4 = (y/4)*4 --> 4x = y	Similar as in step #1, we want to simplify the equation and get closer to the ultimate goal. So we multiple 4 on both sides of the equation.
4x = y --> y = 4x	Following the equation rule of symmetric property, which is if a = b, then b = a, we can put y on the left side of the equation.
y = 4x --> y/x = (4x)/x = 4 --> y/x = 4	Finally, we divide both sides of the equation by x to get the value of y/x.

Note. The multiplication/division rule for equations tell us that every term on both sides of an equation can be multiplied or divided by the same term (except zero) without changing the solution set of the equation. For example, given 0.5x = 4, we can get simplified equation (0.5x) * 2 = 4 * 2, x = 8.

Answer to Problem #4.

Final answer: 3.30 Your answer: ____ Correct: Yes ___ No ____

Steps	**Rules & explanations**
Σ , represents a sum; x_i, represents the *i*th value in the data, in other words, the subscript *i* represents the specific cases and it varies from 1 to 10; n, represents the total number of the group.	Get familiar with statistics/mathematics symbols; Calculate each component in the formula
$\sum x_i$= **1** + **6** + **3** + 0 + 2 + 3 + 4 + 6 + 5 + 3 = (**1** + **6** + **3**) + 0 + **2** + **3** + 4 + 6 + **5** + 3 = 10 + (**2** + **3** + **5**) + **4** + **6** + 3 =10 + 10 + (**4** + **6**) + 3 =10 + 10 + 10 + 3 = 33; n = 10;	Follow the order of operations and do the calculation: Use the table presented in problem #3. First, check if there is any part that has the priority to calculate; the answer is no. Second, do additions from left to right; you may notice that combining a couple of terms, e.g., 1, 6, 3, and making them as 10s can easy your calculation (changing the order of the addition won't change the final results).
Therefore, the average frequency = $\frac{\sum x_i}{n}$ = 33/10 = 3.30	Once the total frequency is ready, you use the total to divide by 10.

Answer to Problem #5.

Final answer: 0.65 Your answer: ___ Correct: Yes ___ No ____

First thing first: Let's review what is the conditional probability. A conditional probability is the probability of one thing (A) being true **given** that another thing (B) is true. It can be written as P (A | B). Pay attention to the assumption that event A and B are dependent. If A and B are independent, there is no need for calculating the conditional probability.

In this problem scenario, event A represents students who pass the 1st quarter exam. P(A) represents the probability that students passed the 1st quarter exam. Event B represents students who pass the 2nd quarter exam. P (B) represents the probability that students passed the 2nd quarter exam. P (A ∩ B) represents the probability that students passed both exams. Finally, P (A | B) represents given those who pass the 2nd quarter exam (event B), the probability of those who also pass the 1st quarter exam.

Steps	**Rules & explanations**
To get the P (A | B), we just need to get P (A ∩ B) and P(B).	Following the definition of conditional probability, you will need to know the P (A ∩ B) and P(A).
P (A ∩ B) = 20% = .26, P(B) = 40% = .4	Fill in those expressions using known information.

P (A | B) = P (A ∩ B) / P(B) = .26/.4 = .26/.4 = 26/40 $=\frac{26}{40}=\frac{2*13}{2*20}=\frac{13}{20}=\frac{13*5}{20*5}=\frac{65}{100}=$.65	Finally, you just need to divide the value P (A ∩ B) by the value P(B).

So, the final answer is .65. It means 65% of students who pass the second quarter exam also pass the first quarter exam.

*You can also illustrate this problem using a contingency table (or crosstab):

There are two events: students took 1st and 2nd exam in their class. There are only four different combinations for the conditional probability: Students who passed 1st exam and 2nd exam, who failed 1st and 2nd exam, who passed 1st exam but failed 2nd exam, and who failed 1st exam but passed 2nd exam (in the table, the proportions/probabilities to represent those four situations are shown as p, q, r, and s).

	Pass 2nd exam	Fail 2nd exam	Marginal
Pass 1st exam	p	r	
Fail 1st exam	s	q	
Marginal	40%		1

Then, go back to the question. In the question stem, we know that p = 26% = .26; and p + s = 40% = .40. The question is what is p / .40? The tricky part is the change of the population:

p is calculated with the whole population; but the question is about the subpopulation (which is among only the 40% of the whole population -that's why this question is related to conditional probability). And because it is a subpopulation, the answer, which is .65 or 65%, is much larger than *p* (.26).

Table A2

Rubrics for Scoring Procedural Statistical Knowledge Test Items

Scores	Scoring rubric	Example
2	Show each *detailed* steps/procedure on the problem-solving process, list the relevant formula/equation right, and answer is correct;	$4^2 + (-3)^2 + 0^2 + (-1)^2 = 16 + 9 + 0 + 1 = 26$
1	*Partially* show the necessary steps/procedure, but the calculation (whole or partial) is incorrect; the result might be correct;	$4^2 + (-3)^2 + 0^2 + (-1)^2 = 16 + (-9) + 0 + (-1) = 6$
0	*No show* of the necessary steps/procedure, calculation is incorrect, but result is incorrect, or irrelevant writings.	$(4+4)^2 + (-3)^2 + 0^2 + (-1)^2 =$

APPENDIX E

Perceived Utility Value of Statistics Questionnaire

Part 1. The Writing Task

Instruction.

Thank you for taking the commitment survey and prior statistical knowledge questionnaires. Now, please type a short essay (approximately 1 – 3 paragraphs in length) briefly describing the potential relevance of statistical knowledge to your own life, or to the life of college students in general. Of course, you may probably need more practice with your statistical knowledge to really appreciate its personal relevance. However, for the purpose of this writing exercise, please focus on how statistical knowledge could be useful to you or to college students in general and give 1 – 2 examples. In general, there is no time limit for writing this essay. As a recommendation, you may use about 10 – 15 minutes. Please type below:

*Reminder: Please type at least one paragraph with at least one specific example (including details such as facts, personal stories, experiences, self-reflection, or other detailed information).

Part 2. Post-section questionnaire

Reflection on your statistics writing task experience.

Now think about what you wrote previously. Rate the following statements regarding your writings. Rethink the statistics topics that you typed in your short essay:

1. How relevant it is to undergraduate students' future career in general?

- Extremely
- Very
- Moderate
- Slightly
- Not at all

2. How relevant it is to your future career?

- Extremely
- Very
- Moderate
- Slightly
- Not at all

3. How important it is to the undergraduate students' daily life in general?

- Extremely
- Very
- Moderate
- Slightly
- Not at all

4. How important it is to your daily life?

- Extremely
- Very
- Moderate
- Slightly
- Not at all

Note. Exploratory factor analysis for individuals' beliefs about "responsibility for perceived utility value of statistics" (Q1-4): One factor (Eigenvalue = 2.95) was comprised of 4 items and explained 74% of variance with factor loading from 0.82 to 0.90. $n = 40$.

Part 3. Remembering the utility value of statistics

3.1 Remembering the utility value of conceptual statistical knowledge.

<u>Instruction</u>. Reflect on your previous experience of doing the tasks in statistical conceptual knowledge questionnaire in your paper-and-pencil section (i.e., the concept map assessment).

For each of following statement, please rate:	Not agree at all			Neutrally			Totally agree
1. In general, it's interesting to work on the assessment of statistical conceptual knowledge.	1	2	3	4	5	6	7
2. I like doing this type of assessment of statistical conceptual knowledge.	1	2	3	4	5	6	7
3. It's worthwhile to me that I do well in mastering statistical conceptual knowledge.	1	2	3	4	5	6	7
4. It's worthwhile to me that being good at solving statistical problems using statistical conceptual knowledge.	1	2	3	4	5	6	7
5. It's important to me that I can do well in this type of assessment in the future.	1	2	3	4	5	6	7
6. It's useful to learn advanced statistical conceptual knowledge for what I want to do in the future.	1	2	3	4	5	6	7
7. It's useful to learn advanced statistical conceptual knowledge for my daily life outside school.	1	2	3	4	5	6	7
8. Comparing to other students, I expect I do better in this type of statistical knowledge assessment.	1	2	3	4	5	6	7
9. I think I will do well in this type of conceptual knowledge assessment if there is one in my statistics course.	1	2	3	4	5	6	7
10. I am very good at this type of statistical conceptual knowledge assessment.	1	2	3	4	5	6	7
11. If I was to order all the students in my statistics class from the worst to the best in this type of statistical knowledge test, I will put me at the top.	1	2	3	4	5	6	7
12. In general, I have mastered a lot of statistical conceptual knowledge earlier than this semester's statistics course.	1	2	3	4	5	6	7

3.2 Remembering the utility value of procedural statistical knowledge.

Instruction. Reflect on your previous experience of solving the problems in procedural statistical knowledge questionnaire in your paper-and-pencil section.

For each of following statement, please rate:	Not agree at all			Neutrally			Totally agree
1. In general, it's interesting to work on the assessment of statistical procedural knowledge.	1	2	3	4	5	6	7
2. I like doing this type of assessment of statistical procedural knowledge.	1	2	3	4	5	6	7
3. It's worthwhile to me that I do well in mastering statistical procedural knowledge.	1	2	3	4	5	6	7
4. It's worthwhile to me that being good at solving statistical problems using statistical procedural knowledge.	1	2	3	4	5	6	7
5. It's important to me that I can do well in this type of assessment in the future.	1	2	3	4	5	6	7
6. It's useful to learn advanced statistical procedural knowledge for what I want to do in the future.	1	2	3	4	5	6	7
7. It's useful to learn advanced statistical procedural knowledge for my daily life outside school.	1	2	3	4	5	6	7
8. Comparing to other students, I expect I do better in this type of statistical knowledge assessment.	1	2	3	4	5	6	7
9. I think I will do well in this type of procedural knowledge assessment if there is one in my statistics course.	1	2	3	4	5	6	7
10. I am very good at this type of statistical procedural knowledge assessment.	1	2	3	4	5	6	7
11. If I was to order all the students in my statistics class from the worst to the best in this type of statistical knowledge test, I will put me at the top.	1	2	3	4	5	6	7
12. In general, I have mastered a lot of statistical procedural knowledge earlier than this semester's statistics course.	1	2	3	4	5	6	7

APPENDIX F

Demographics and Academic Information

Demographic Information.

1. How old are you? Please specify (only include years, e.g., 20): ______

2. What is your gender?

- Male
- Female
- Prefer not to say
- Prefer to self-describe: _____

3. How do you report your ethnicity on official documents?

- American Indian/Alaskan Native
- Arab/Arab American
- Asian/Asian American

  Check if apply: East Asian (e.g., Chinese, Korean, Japanese, etc.)

  Check if apply: South Asian (e.g., India, Pakistan, Bangladesh, Nepal)

  Check if apply: South East Asian (e.g., Vietnam, Thailand, Indonesia)

  Asian/Asian American: Other (Please specify) _____

- Black/African American

  Black/Other Country of Origin (Please specify) _____

- Mexican/Mexican American

  Latino/Other Country of Origin (Please specify) _____

- Native Hawaiian/Pacific Islander

  Native Hawaiian/Pacific Islander (Please specify) _____

- White/Caucasian
- Other (Please specify) _____

4. How long have you (and/or your family) lived in the United States? (Check one that apply)

- I am (we are) new to this country (I was/We were not born here).
- My parents first came here.
- My grandparents first came here.
- My family has been here for many generations.

5. Are you first generation of college students in your family?

- Yes
- No
- Not Sure/Don't Know

Academic Information.

6. In the following questions 6(1) - 6(3), please write down your answers:

Q6(1). Number of credit hours have earned toward the degree that you are seeking at UIC (Please specify; this semester not included)

Q6(2). Number of high school mathematics and statistics course(s) completed (Please specify)

Q6(3). Number of undergraduate mathematics and statistics course(s) completed at UIC (Please specify; do not include this semester)

7. What mathematics course(s) have you completed in your high school? (Check all that apply)

- Algebra I
- Geometry
- Algebra II
- Pre-Calculus
- Calculus
- Trigonometry
- Statistics
- Other (Please specify) ____

8. What is your (intended) major at UIC?

- Art/Humanities
- Business
- Education
- Engineering
- Medicine/Pre-Medicine
- Nursing/Pre-Nursing
- Law/Pre-Law
- Natural sciences
- Social sciences (e.g, psychology, sociology)
- Statistics/Mathematics
- Not Decided/Don't know
- Other (Please specify) ____

9. Are you enrolled in any minor degree's program at the current university? If yes, please specify; Otherwise, please respond "No". _______

10. Have you transferred to your current university from another college/university?

- Yes
- No

11. What year are you in your field of study at your current university? (If you have transferred here, only count your current university years.)

- First year
- Second year
- Third year
- Fourth year
- Other (Please specify) ____

12. On a 0 – 4 scale, what is your current GPA (not include this semester)? Please give a single number with two decimals, e.g., 3.15. _______

13. Do you receive any academic accommodation (e.g., individualized education program) currently at your current university?

- Yes
- No

14. To prepare for studying statistics, have you ever used any online resources, e.g., massive open online courses (MOOC), blogs, or youtube.com? (Check one)

- Yes (please specify what resource you have used) _____
- No

Note. Collected responses from questions 4, 5, 9, 10, 13, and 14 were not included in the final analysis since they did not directly relate to the research questions in my dissertation.

APPENDIX G

Table A1

Results of t-tests on Statistics Exam Scores: Stayed-in versus Dropped-out Participants in their Post-Test Survey

	Stayed-in *M* (*SD*) (*n* = 101)	Dropped-out *M* (*SD*) (*n* = 17)
1st quarter exam	78.52 (14.68)	77.41 (13.58)
Midterm exam	75.13 (16.35)	68 (18.11)
Final exam	70.45 (18.83)	61.59 (23.39)

Note. Two-tailed *t*-tests were conducted and there was no significant difference. Significance level: *, $p < .05$; **, $p < .01$; ***, $p < .001$ with Bonferroni corrections. Results from Levene's test for equality of variances showed no violation on the assumption of equal variance between groups.

Table A2

Results of t-tests on Statistics Exam Scores: Stayed-in versus Dropped-out Participants in their Delayed Post-Test Survey

	Stayed-in *M* (*SD*) (*n* = 84)	Dropped-out *M* (*SD*) (*n* = 34)
1st quarter exam	77.64 (14.43)	79.75 (14.55)
Midterm exam	73.93 (17.67)	74.00 (14.81)
Final exam	69.32 (19.03)	68.81 (21.51)

Note. Two-tailed *t*-tests were conducted and there was no significant difference.

VITA

NAME: **Kuan Xing**

TITLE: **Director of Assessment and Research,** Center for Healthcare Improvement and Patient Simulation (CHIPS), University of Tennessee Health Science Center

EDUCATION: **B.E.** in Information Engineering, China University of Mining and Technology, China 2006

M.Ed. in Developmental and Educational Psychology, Anhui Normal University, China 2010

M.Ed. in Measurement, Evaluation, Statistics, and Assessment, University of Illinois at Chicago 2015

Ph.D. in Educational Psychology, University of Illinois at Chicago 2020

PUBLICATIONS: Park, Y. S., & **Xing, K**. (2019). Rater model using signal detection theory for latent differential rater functioning. *Multivariate Behavioral Research*, *54*, 492-504.

Jani, P., Blood, A. D., Park, Y. S., **Xing, K**., & Mitchell, D. (published online). Simulation-based curricula for enhanced retention of pediatric resuscitation skills: A randomized controlled study. *Pediatric Emergency Care*.

Park, Y. S., **Xing, K**., & Lee, Y.-S. (2018). Explanatory cognitive diagnostic models: Incorporating latent and observed predictors. *Applied Psychological Measurement*, *42*, 376-392.

Thorkildsen, T. A., & **Xing, K**. (2016). Facebook as a tool for enhancing communication and self-expression. In S. Tettegah, & G. Kien (Eds.), *Emotions, technology & social media* (pp. 117-138). San Diego, CA: Elsevier.

Park, Y. S., Lee, Y.-S., & **Xing, K**. (2016). Investigating the impact of item parameter drift for Item Response Theory models with mixture distributions. *Frontiers in Psychology*, *7*:255.

Park, Y. S., Hurm, M., **Xing, K**., Anderson, L., Kamin, C., & Yudkowsky, R. (2014). Investigating the psychometric properties of scoring based on live-interaction and videotaped observations: Focusing on communication and interpersonal skills. *Simulations in Healthcare*, *9*, 209.

Park, Y. S., Hyderi, A., Bordage, G., **Xing, K**., & Yudkowsky, R. (2016). Inter-rater reliability and generalizability of patient note scores using a scoring rubric based on the USMLE step-2 CS format. *Advances in Health Sciences Education, 21*, 761-773.

Park, Y. S., Lineberry, M, Hyderi, A, Bordage, G., **Xing, K**., & Yudkowsky, R. (2016). Differential weighting for subcomponent measures of integrated clinical encounter scores based on the USMLE Step-2 CS examination: Effects on composite score reliability and pass-fail decisions. *Academic Medicine (suppl.)*, *91*, S24-S30.

Xing, K., Chico, E., Lambouths, D., Brittian, A. S., & Schwartz, S. J. (2015). Identity development in adolescence: Implications for youth policy and practice. In E. P. Bowers, G. J. Geldhof, S. Johnson, L. Hilliard, R. Hershberg, J. V. Lerner, & R. M. Lerner (Eds.), *Promoting healthy development for America's youth: Lessons learned from the 4-H study of positive youth development* (pp. 187-208). Springer.

Leathers, S. J., Vande Voort, B. L., **Xing, K**., Walsh, K., Spielfogel, J. E., Annes, L., Frizzell, T., & Fleary-Simmons, D. (2019). Countdown to 21: Outcomes from a transition support program for older youth exiting foster care. *Child Welfare, 97(6)*, 53-76.

PAPER PRESENTATIONS: **Xing, K**., Dai, T., Kaplan, A., Perez, T., Cromley, J. G., Balsai, M. J., & Mara, K. R. (August, 2019). *Self-perceived conscientiousness on introductory biology learning: For whom does it matter?* Poster will be presented at the APA 2019 Convention, Chicago, IL.

Xing, K., Lin, Q., & Park, Y. S. (April, 2019). *Multilevel analysis incorporating multiple covariates for independent and higher-order cognitive diagnosis models*. Paper presented at 2019 National Council on Measurement in Education (NCME) conference, Toronto, CA.

Xing, K., & Becker, K. (April, 2017). *Quantifying the item order effect on item difficulty in large-scale testing.* Paper presented at the 2017 NCME annual meeting, San Antonio, TX.

Xing, K., Park, Y. S., & Thorkildsen, T. A. (April, 2016). *Regressing multiple predictors into a cognitive diagnostic model*. Poster presented at the 2016 NCME annual meeting, Washington, D.C.

Park, Y. S., Lee. Y.-S., **Xing, K**., & Lim, M. (April, 2016). *Incorporating latent and observed predictors in cognitive diagnostic models*. Paper presented at the 2016 NCME annual meeting, Washington, D.C.

Xing, K., Park, Y. S., & Thorkildsen, T. A. (July, 2015). *Modeling online social networking behaviors: A multidimensional mixture IRT approach*. Poster presented at the 79th Annual Meeting of Psychometric Society, Beijing, China.

Park, Y. S., Lee, Y.-S., & **Xing, K**. (July, 2015). *Mixture higher-order diagnostic classification model with covariates*. Paper presented at the 79th Annual Meeting of Psychometric Society, Beijing, China.

Xing, K., Park, Y. S., & Thorkildsen, T. A. (April, 2015). *Application of mixture IRT to online social networking behaviors*. Poster presented at the 2015 NCME Annual Meeting, Chicago, IL.

Park, Y. S., Kim, M. H., & **Xing, K**. (April, 2015). *Effects of scoring designs on rater precision and classification*. Individual paper presentation at the 2015 NCME Annual Meeting, Chicago, IL.

Park, Y. S., & **Xing**, **K**. (April, 2014). *Evaluating and adjusting for rater effects in standard setting using the Angoff method*. Paper presented at the 2014 AERA Annual Meeting, Philadelphia, PA.

Thorkildsen, T. A., & **Xing, K**. (August, 2013). *Relations between social capital and achievement among Facebook user*. Poster presented at the 2013 APA Convention, Honolulu, HI.

HONORS: 2018 **Dissertation Research Papers Fund**, *Office of Research*, *College of Education, University of Illinois at Chicago*

2013 **APA Student Travel Award**, *American Psychological Association*

PROFESSIONAL MEMBERSHIP: Society for Simulation in Healthcare,

American Educational Research Association,

National Council on Measurement in Education,

American Psychological Association