THE EFFECT OF MOTIVATION AND ATTITUDE TOWARDS STATISTICS ON CONCEPTUAL UNDERSTANDING OF STATISTICS

by

PAUL JAMES HOLMES

(Under the Direction of Seock-Ho Kim)

ABSTRACT

This study explored relationships between a student's motivation for attending college, his or her attitude towards statistics, and the conceptual understanding he or she derived from an introductory statistics course. The purpose of this study was to determine the appropriate dimensionality of the Academic Motivation Scale (AMS) and to use that scale, as well as the Survey of Attitudes towards Statistics (SATS-28) to predict conceptual understanding of statistics, measured by the Comprehensive Assessment of Outcomes for a first course in Statistics (CAOS). This was done through the use of SEM and stepwise multiple regression. Finally, the differences in conceptual understanding between students at opposite extremes on the various dimensions of motivation were compared.

The study sample consisted of 1350 students taking an introductory statistics course at the University of Georgia in Fall 2012. Major findings confirm the four dimension structure, consisting of *Intrinsic Motivation*, *Amotivation*, *External Regulation*, and *Identified Regulation*, discussed in Smith et al. (2012). From these four motivational factors, and the four attitudinal factors of *Affect*, *Cognitive Competence*, *Value*, and *Difficulty*, the study determined that the most significant factors in predicting conceptual understanding were (1) the student's identified regulation (representing the amount that the student recognizes the behavior as important and performs it out of choice); (2) amotivation (the lack of motivation to attend college, typically due to a feeling of not being able to succeed); (3) the extent to which the student's reasons for going to college are externally regulated (such as career goals); (4) the amount of cognitive competence the student has in their ability of do well in statistics, and; (5) the value and perceived relevance that the student places on statistics in his/her own life.

INDEX WORDS: Motivation, Statistics attitudes, Statistics anxiety, Conceptual understanding

THE EFFECT OF MOTIVATION AND ATTITUDE TOWARDS STATISTICS ON CONCEPTUAL UNDERSTANDING OF STATISTICS

by

PAUL JAMES HOLMES

BS, University of Liverpool, England, 1997

MS, University of Georgia, 1999

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2014

THE EFFECT OF MOTIVATION AND ATTITUDE TOWARDS STATISTICS ON

CONCEPTUAL UNDERSTANDING OF STATISTICS

by

PAUL JAMES HOLMES

Major Professor: Seock-Ho Kim

Committee: Jaxk Reeves
Stephen Cramer
Laura Lu

Electronic Version Approved:

Julie Coffield
Interim Dean of the Graduate School
The University of Georgia
August 2014

ACKNOWLEDGEMENTS

I would like to thank my Major Professor and dissertation committee chair, Dr. Seock-Ho Kim. Despite numerous months where my teaching schedule would see me go off-the-grid, he was always ready to help whenever I popped my head around his office door with a stack of questions. Similarly, committee members Drs. Jaxk Reeves, Stephen Cramer and Laura Lu, who showed great patience with a student that they may at times have felt was destined to spend a lifetime in academia without ever graduating. Thanks also go to Dr. Stephen Olejnik, my former committee chair, who retired before I completed this process but was incredibly helpful in helping me formulate research ideas. Additionally, for their help in the collection of the data, I would like to thank Jack Morse and Dr. Jennifer Kaplan.

Huge amounts of gratitude also go to my wife, Sandra, who never let me quit this process (even though she secretly knows that if I have a PhD then she’ll want one too) and my parents, Maureen and Les Holmes, who will tell literally everyone they know that they were mentioned in my dissertation (if I ever let them read it). Finally, after twelve years of false dawns and halted progress, possibly the biggest motivation in helping me finish my PhD has been my wonderful children, Liam and Sophia. I’ll be the happiest dad in the world when you guys come to my graduation.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS iv

LIST OF TABLES vii

LIST OF FIGURES ix

CHAPTER

1 INTRODUCTION 1

Overview 1

2 REVIEW OF THE LITERATURE 3

Statistical Understanding and Common Misconceptions 3

The Role of Motivation in Conceptual Understanding 8

The Role of Statistics Attitudes in Conceptual Understanding 14

3 METHODS 17

Instruments Used 17

Research Questions 18

Data Collection 19

4 ANALYSIS 27

Analysis 1. Principal Component Analysis to Determine Dimensionality of Academic Motivation Scale (AMS) 29

Analysis 2. Construction of SEM Model to Explain Conceptual Understanding of Statistics 34

Analysis 3. Construction of Multiple Regression Models to Explain Conceptual Understanding of Statistics 38
Analysis 4. Item Analysis to Compare Conceptual Understanding of Students with Different Extremes of Motivational Traits 51
5 SUMMARY AND DISCUSSION 61
Summary of the Study 61
Limitations and Implications for Future Research 68
Teaching Implications 69
Conclusion 70
REFERENCES 72
APPENDICES
A 40 Measured Learning Outcomes of CAOS Test Items 81
B 28 Items of the SATS-28 Survey 85
C Demographic Survey 87
D 28 Items of the AMS 88
E Mplus Program for SEM Model 90
F Correlations between CAOS 4 Score, Motivation and Attitude Subscales, and Demographic Variables 91

LIST OF TABLES

Page

Table 1: Summary of Survey Instruments Considered ... 22

Table 2: Summary of Survey Instruments Used ... 23

Table 3: Summary of Survey Instrument Completion ... 24

Table 4: Demographic Breakdown of Sampled Students ... 25

Table 5: Summary Statistics for Academic Motivation Scale (AMS) Items ... 26

Table 6: Summary Statistics for Student Attitudes towards Statistics (SATS-28) Items ... 27

Table 7: Number and Percentage of Students Answering each CAOS 4 Item Correctly ... 28

Table 8: Summary Statistics for Overall CAOS 4 Scores ... 28

Table 9: Factor Loadings for Academic Motivation Scale Items ... 31

Table 10: Eigenvalues and Percentage of Variance Explained by Each Factor, along with Summary Statistics and Consistency Measures for Included Items ... 33

Table 11: Summary Statistics for AMS and SATS-28 Subscales and Demographic Variables ... 39

Table 12: Table of Parameter Estimates for Stepwise Regression Model Predicting CAOS Score from Motivation and Statistics Attitudes Subscales ... 40

Table 13: Summary of Stepwise Selection Process for Predicting CAOS Score from Motivation and Statistics Attitudes Subscales ... 41

Table 14: Table of Parameter Estimates for Stepwise Regression Model Predicting CAOS Score from Motivation and Statistics Attitudes Subscales and Interactions ... 42

Table 15: Summary of Stepwise Selection Process for Predicting CAOS Score from Motivation and Statistics Attitudes Subscales and Interactions ... 43

Table 16: Table of Parameter Estimates for Final Regression Model Predicting CAOS Score from Motivation and Statistics Attitudes Subscales and Interactions ... 44

Table 17: Table of Parameter Estimates for Stepwise Regression Model Predicting CAOS Score from Motivation and Statistics Attitudes Subscales and Demographic Characteristics....45

Table 18: Summary of Stepwise Selection Process for Predicting CAOS Score from Motivation and Statistics Attitudes Subscales and Demographic Characteristics....46

Table 19: Table of Parameter Estimates for Stepwise Regression Model Predicting CAOS Score from Motivation and Statistics Attitudes Subscales and Interactions, as well as Demographic Characteristics....48

Table 20: Summary of Stepwise Selection Process for Predicting CAOS Score from Motivation and Statistics Attitudes Subscales and Interactions, as well as Demographic Characteristics....49

Table 21: Table of Parameter Estimates for Final Regression Model Predicting CAOS Score from Motivation and Statistics Attitudes Subscales and Interactions, as well as Demographic Characteristics....50

Table 22: Comparison of CAOS Item Performance between Low and High Intrinsically Motivated Students....56

Table 23: Comparison of CAOS Item Performance between Low and High Amotivation Students....57

Table 24: Comparison of CAOS Item Performance between Low and High External Regulation Students....58

Table 25: Comparison of CAOS Item Performance between Low and High Identified Regulation Students....59

LIST OF FIGURES

Page

Figure 1: Structural equations model of college motivation and statistics attitude subscales and conceptual understanding of statistics..35

Figure 2: Arrangement of and correlations between motivation subscales consistent with Ryan & Deci (2000) taxonomy of human motivations..63

Figure 3: Correlation between statistics attitude subscales..63

Figure 4: Correlation between subscales and CAOS scores...64

CHAPTER 1

INTRODUCTION

Overview

In an information-rich society, the need for statistical literacy in the consumers of such information is vital. Enrollment in Advanced Placement Statistics courses has increased more than twentyfold since 1997, the first year the exam was offered, to 152,750 in 2012 (Rodriguez, 2012) with enrollment in AP Statistics classes increasing more rapidly than in any other A.P. area (Boslaugh & Watters, 2008). The Guidelines for Assessment and Instruction in Statistics Education (GAISE) Project also reported that, whereas enrollment in statistics courses was 27% of the size of enrollment in calculus courses in 1970, 30 years later that ratio had increased to 74% (GAISE, 2006). Yet after completing these classes many students are left with only a rudimentary, misconception-laden understanding of statistics. Additionally, since many of these students do not necessarily see the relevancy of statistics to their major, taking a statistics course is often seen as a negative experience, with a student's lone motivation being to achieve a passing grade, rather than gaining statistical competency.

The GAISE (2006) report noted that statistics students now "take statistics earlier in their lives, increasingly often in high school; few are drawn to statistics by immediate practical need; and there is great variety in their levels of quantitative sophistication. As a result, today's teachers face challenges of motivation and exposition that are substantially greater than those of a half-century ago."

The aim of this study is to address the role that the different motivations and experiences that students bring into an introductory statistics course have on their ability to understand statistics, and whether students with differing motivations and attitudes towards statistics develop a more accurate conceptual understanding with less misconceptions.

CHAPTER 2

REVIEW OF THE LITERATURE

Statistical Understanding and Common Misconceptions

In examining what students should be expected to know about – and do with – statistics in elementary and middle grades, Friel et al. (1997) suggested that students should develop a 'data sense' that involves, "being comfortable with posing questions, collecting and analyzing data, and interpreting the results in ways that respond to the original question asked. It also includes comfort and competence in reading, listening to, and evaluating reports based in statistics, such as those found in newspapers, magazines, television, and other forms of popular press." To develop this 'data sense', the National Research Council (1991) suggested, "what is needed is a variety of activities, including discussion among pupils, practical work, practice of important techniques, problem solving, application to everyday situations, investigational work, and exposition by the teacher."

Additionally, in a list of recommendations for the teaching of introductory statistics, Cobb (1992) made several recommendations that formed the basis of the GAISE report. This report recommended the emphasis of statistical literacy and the development of statistical thinking, the use of real data, a stressing of conceptual understanding over procedural knowledge, an active learning approach to teaching, the use of technology for both analysis and for garnering conceptual understanding, and assessments that both improve and evaluate student learning. These recommendations ushered in a reform movement in statistics education that saw a much heavier focus on understanding of statistical concepts and analysis of data in many introductory statistics courses. This topic was discussed in detail by Moore (1997).

Much research focuses on the development of *statistical literacy*, *statistical reasoning*, and *statistical thinking*, yet the delineation between these topics is often unclear. Chance (2002) notes that the terms are often used interchangeably to represent a desired learning goal of conceptual understanding, often to distinguish said goal from the number crunching and formula-based approach prevalent in much statistics instruction. In their Teaching Statistical Concepts textbook, Hawkins, Jolliffe, and Glickman (1992) also discuss statistical reasoning and thinking together, without distinguishing between the two. delMas (2002) argued that these three learning outcomes, while distinct, do have some overlap – with statistical literacy being the foundation for the other two learning outcomes.

Statistical literacy has variously been described as the ability to interpret and evaluate statistical information and the ability to discuss such information (Gal, 2000). Garfield (1999) described it as an understanding of statistical language and terminology and an ability to interpret graphs and tables and to read and make sense of statistics in the news and media, whereas Schield (1999) describes it as the ability to use statistics as evidence in arguments and to think critically about statistics. In noting the broadness of these various definitions, Rumsey (2002) partitions the idea of statistical literacy into two distinct learning goals – statistical competence (the basic knowledge involved in statistical reasoning) and statistical citizenship (the ability to make sense of the wealth of data experienced in everyday life).

Statistical reasoning can be defined as “the way people reason with statistical ideas and make sense of statistical information. Underlying this reasoning is a conceptual understanding of important ideas, such as distribution, center, spread, association, uncertainty, randomness and sampling" (Garfield, 2002). Lovett (2001) described it as, "the use of statistical tools and concepts (e.g., hypothesis testing, variation, correlation) to summarize, make predictions about, and draw conclusions from data."

The goal of developing *statistical thinking* was discussed by Chance (2002), who distinguished it from statistical literacy and reasoning in the following way: "Perhaps what is unique to statistical thinking, beyond reasoning and literacy, is the ability to see the process as a whole (with iteration), including "why," to understand the relationship and meaning of variation in this process, to have the ability to explore data in ways beyond what has been prescribed in texts, and to generate new questions beyond those asked by the principal investigator. While literacy can be narrowly viewed as understanding and interpreting statistical information presented, for example in the media, and reasoning can be narrowly viewed as working through the tools and concepts learned in the course, the statistical thinker is able to move beyond what is taught in the course, to spontaneously question and investigate the issues and data involved in a specific context."

Since the concepts of statistical literacy, reasoning, and thinking are so often confused and used almost interchangeably, delMas (2002) suggests key words that ask students to demonstrate abilities in a particular domain. Key indicators of assessments of statistical literacy include *identify*, *describe*, *re-phrase*, *translate*, *interpret*, and *read*; measures of reasoning often ask the student *why*, *how* or to *explain* a process; whereas identifiers of statistical thinking assessments often include the terms *apply*, *critique*, *evaluate*, or *generalize*.

Much research on statistical understanding focuses on student misconceptions. Utts (2003) identified seven statistical topics that are often misunderstood by citizens, journalists, and even some researchers as areas to focus on in order to improve statistical literacy. Those areas are: when a cause and effect relationship can be concluded; the difference between statistical significance and practical importance; the difference between ‘no effect’ and ‘no statistically significant effect’; common sources of bias; that coincidences and seemingly improbable events are not uncommon; “confusion of the inverse” with conditional probability; and misunderstanding of the nature of variability. Other misconceptions of statistics students were identified by Garfield (2002), including misconceptions about averages, a belief that the ratio of the sample size to the population size needs to be high to have a “good” sample, and the gamblers’ fallacy (that a particular outcome is “due” if it hasn’t occurred in a while). Misconceptions about probability (Hirsch & O’Donnell, 2001), statistical inference (Sotos et al., 2007), confidence intervals (Kalinowski, 2010) and hypothesis testing (Sotos et al., 2009) have also been researched. In looking at student misconceptions in Statistics, Konold (1995) found that many of these misconceptions persevere, as students incorporate and adjust new knowledge to fit into their previously held (incorrect) beliefs.

Whereas the occurrence of these misconceptions in some students may not be surprising, evidence suggests that even high-performing students in statistics courses may suffer from many of the same misinterpretations. Matthews and Clark (2003) examined the statistical understanding of a selection of high-scoring statistics students (received an ‘A’ grade in an elementary probability and statistics class) several weeks after completion of the course and found a consistently poor understanding of means, standard deviations and the central limit theorem, containing a “heavy reliance (both conscious and unconscious) on algorithmic procedures which are not understood.” Garfield (1998) also noted that students who achieve good grades in a statistics course often did poorly on measures of statistical reasoning. These disappointing results from high-scoring students suggest that, rather than measuring a student’s statistical literacy or ability to apply statistical reasoning, many measures of performance in statistics courses (be they tests, homework or other coursework) instead tend to over-emphasize computation and procedure. Konold (1995) warns however, that "in trying to assess competencies at the most rudimentary level, or to not overwhelm students with problems that require legitimate problem solving and reasoning, the teacher inadvertently leads students to believe that routine skills and memorized formulae are the important stuff."

Many attempts to measure statistical literacy and reasoning with instruments measuring conceptual understanding rather than performance of rote procedures are limited in scope. The Statistics Concepts Inventory (SCI, Stone et al., 2003) is a multiple choice assessment tool designed to measure statistical understanding in engineering students, yet it may not be applicable to other majors (Reed-Rhoads et al., 2006). The Statistics Reasoning Assessment (SRA, Garfield, 1998) was designed to measure reasoning in statistics and probability, with items testing for the presence on many of the common previously discussed statistics misconceptions, yet the scope of the SRA was limited by the absence of items on statistical inference (Garfield, 2003).

The Comprehensive Assessment of Outcomes in Statistics (CAOS, Garfield et al., 2002) test does not appear to suffer from any such limitations. CAOS was developed by the Assessment Resource Tools for Improving Statistical Thinking (ARTIST) project. ARTIST was funded by the National Science Foundation (NSF) as a means to develop reliable and valid assessment items to measure students' statistical literacy, statistical reasoning, and statistical thinking. The ARTIST website (https://apps3.cehd.umn.edu/artist/caos.html) provides a range of assessment items for teaching first courses in statistics, including a selection of online multiple choice tests on various topics, as well as the latest version (CAOS 4) of the CAOS test, a 40-item test of overall conceptual statistical understanding. The statistical topics covered in these 40 questions are shown in Appendix A.

The development, reliability and validity of the CAOS test were addressed by delMas et al. (2007). The 40 items on the CAOS 4 post-test produced an alpha coefficient of .77, based on a sample of 10,287 students. Additionally, 18 expert raters unanimously agreed that CAOS 4 measures important basic learning outcomes, with 94% agreement on it measuring important learning outcomes. All raters agreed that, "CAOS measures outcomes for which I would be disappointed if they were not achieved by students who succeed in my statistics courses." Whilst some raters suggested additional topics that were missing from the scale, there was no agreement on those topics, suggesting that CAOS is a valid measure of important learning outcomes in an introductory statistics course.

The Role of Motivation in Conceptual Understanding

The role that motivation plays in development of conceptual understanding has not been thoroughly explored in the field of statistics education, but has been examined in other fields. In a study of conceptual understanding of physics, Linnenbrink and Pintrich (2002) found that students adopting mastery goals – that is, learning the topic with the specific goal of mastering it, rather than just to pass a test – saw higher levels of conceptual change than those students adopting performance goals. A study by Pantziara and Philippou (2006) found that sixth-grade students with a strong conceptual understanding of fractions in mathematics had less fear of failure, more self-efficacy and were more likely to have mastery goals than students with lower levels of conceptual understanding. Patrick et al. (2001) also noted that students with mastery goals often had feelings of efficacy, achievement and interest whereas a performance goal orientation was associated with maladaptive achievement beliefs such as low achievement and fear of failure.

Deci and Ryan’s (1985) Self-Determination Theory (SDT) provides a comprehensive, multidimensional model of motivation. This theory suggests that motivation for a specific behavior is regulated by either interior choice or exterior force. When behavior is self-determined, it is considered to have an internal locus of causality, whereas controlled behavior is considered to have an external locus of causality. SDT divides motivation into three categories - intrinsic motivation, extrinsic motivation and amotivation – which can be ordered along a continuum of self-determination or autonomy, with intrinsic motivation considered the most self-determined and amotivation the least.

Komarraju et al. (2009) noted that, “intrinsically motivated individuals tend to have an internal locus of control, are driven to accomplish, seek intellectual stimulation, and are enthusiastic about learning new things. Extrinsically motivated individuals pursue education to achieve contingent goals, rather than for an intrinsic enjoyment of learning.” Students that are intrinsically motivated tend to have mastery goals whereas those that are extrinsically motivated are more likely to have performance goals.

Deci and Ryan (2000) propose that there are four types of **extrinsic motivation** which can also be ordered along an autonomy continuum: *External regulation*, corresponding to the regulation of behavior to achieve positive outcomes or avoid negative ones, represents the lowest level of autonomy. *Introjected regulation* represents the next level of autonomy, whereby the control of behavior is enforced by external sources but the consequences for fulfillment or non-fulfillment are imposed by the individual themselves – such as feelings of guilt, shame, or pride. The third type of extrinsic motivation is *identified regulation*, where the individual identifies the behavior as important and therefore performs it out of choice. With *integrated regulation*, the individual not only recognizes the importance of the behavior and performs it out of choice, but also integrates those identifications with other aspects of their identity. Thus, a student taking a course hoping to achieve a good grade and a student taking the same course because he believes it will be beneficial for his future career are both extrinsically motivated, but the two students differ in their autonomy, with the grade-focused student exhibiting external regulation whilst the career-focused student is using integrated regulation.

Vallerand and O'Connor (1989) propose that there are also three types of **intrinsic motivation**. *Intrinsic motivation to know* can be described as, "the fact of performing an activity for the pleasure and the satisfaction that one experiences while learning, exploring, or trying to understand something new" (Vallerand et al., 1992). *Intrinsic motivation to accomplish things* occurs when the individual engages in a task for the pleasure derived from attempting to accomplish or create something. *Intrinsic motivation to experience stimulation* is the third type of intrinsic motivation, and is demonstrated when an individual engages in a task for stimulation, such as attending a political science class for the excitement of an interesting class discussion or a physical education class for the stimulation of physical activity. These three types of intrinsic motivation have been found to be highly correlated (Pelletier et al., 1995), indicating that they may just be different aspects of one global construct and may have no logical ordering along a motivational continuum (Smith et al., 2012).

A further type of motivation is **amotivation**, which Smith et al. (2010) define as, "the absence of motivation to perform an activity, due to its lack of value to a person, or that person's feeling of incompetence or inability to obtain a desired outcome." Since amotivational behaviors are unregulated and not thought to be controlled by the student, Fortier, Vallerand, and Guay

(1995) argue that it is similar to the idea of "learned helplessness" (Abramson, Seligman, & Teasdale, 1978).

Benware and Deci (1984) performed an experiment comparing a group of intrinsically motivated students told to learn material for the purpose of teaching it to other students with a group of extrinsically motivated students, who were told they would instead be tested on the material. They found that the intrinsically motivated students who learned to teach gave significantly higher ratings for their enjoyment of the material, their interest in the material and their willingness to return. Additionally, whereas those who learned to teach evidenced significantly higher conceptual learning than those who learned to be tested, the two groups did not significantly differ on rote learning, or on time dedicated to learning the material. Pintrich and DeGroot (1990) found that higher levels of intrinsic value were associated with higher levels of student achievement across all performance measures, whereas Gottfried (1990) noted that intrinsically motivated children were more likely to develop strong conceptual understanding, as well as improved memory and high overall achievement in school.

The Motivated Strategies for Learning Questionnaire (MSLQ, Pintrich & DeGroot 1990) is an 81-item survey measuring both motivation and learning. The Motivation section include 31 items measuring goal orientation - both intrinsic and extrinsic, task value, control beliefs about learning, self-efficacy, and test anxiety. The Learning section includes 50 items measuring use of cognitive and metacognitive strategies and students' management of different resources. With only four questions each addressing intrinsic and extrinsic motivation, and with the myriad of factors involved in those areas, these sections would not be able to adequately measure student classroom goals (Shia, 1998). Additionally, some of the MSLQ's sub-scales did suffer from relatively low internal reliability values (Artino, 2005)

The Academic Motivation Scale (AMS, Vallerand et al., 1992) is a 28-item scale that was translated into English from the French-language Echelle de Motivation en Education (EME Vallerand et al., 1989) using appropriate methodological procedures. The scale has seven subscales measuring the three types of intrinsic motivation (intrinsic motivation to know, to accomplish things, and to experience stimulation), three types of extrinsic motivation (external, introjected, and identified regulation, but not specifically integrated regulation) as well as amotivation. Each item on the scale is a response to the question, “Why do you go to college?” with respondents choosing a value from 1 (does not correspond at all) to 7 (corresponds exactly) on items such as, “because with only a high-school degree I would not find a high-paying job later on” and “because I experience pleasure and satisfaction while learning new things.” The English version of the 28 questions of the AMS (along with their subscale identifications) is displayed in Appendix D.

Vallerand et al. (1992) reported that the AMS, “has satisfactory levels of internal consistency (mean alpha value = .81) and temporal stability over a one-month period (mean test-retest correlation = .79). In addition, results of a confirmatory factor analysis (LISREL) confirmed the seven-factor structure of the AMS. Finally, gender differences obtained with the EME were basically replicated with the AMS. In sum, the present findings provide adequate support for the factorial validity and reliability of the AMS and support its use in educational research on motivation.” Cokley et al. (2001) found additional support for the internal consistency of scores from the AMS in a sample of American college students, with alpha coefficients ranging from .70 to .86. These results were consistent with the results of Vallerand et al. (1992).

In this first study of the dimensionality of the AMS with an American college sample however, Cokley et al. (2001) found that the hypothesized seven-factor structure did not have adequate model fit, yet did out-perform other one-factor (with academic motivation as a one-dimensional construct), two-factor (with intrinsic and extrinsic motivation), three-factor (adding amotivation to the two-factor model), and five-factor models (splitting extrinsic motivation into three separate types). The study ultimately added partial support to the construct validity of the AMS, adding that the "seven-factor structure seems to be accurate, but the fact that intrinsic dimensions were not related at all to external academic achievement is counterintuitive and problematic."

Cokley (2000) also concluded that intrinsic and extrinsic motivation, as measured by the AMS, may not be as distinct as Self-Determination Theory suggests. This finding was also supported by Fairchild et al. (2005), who found that Intrinsic Motivation to Know (IMTK) and Intrinsic Motivation to Accomplish (IMTA) shared a correlation of .87, noting that, "with this high correlation, it may be difficult to argue that these different entities of intrinsic motivation represent different subtypes. However, correlational analyses revealed differential relationships between the intrinsic measures and other criteria, suggesting some utility for their separation or continued investigation of their distinctiveness. Additionally, the CFA supported the seven-factor model and it fit better than a model unifying the three types of intrinsic motivation."

Smith et al. (2012) suggested an alternative four-factor configuration of the AMS, consisting of *amotivation*, *external regulation*, *identified regulation*, and *intrinsic motivation*. This model omitted 10 of the 28 AMS items that cross-loaded on multiple factors.

The Role of Statistics Attitudes in Conceptual Understanding

The role of statistics anxiety and statistics attitudes in a student's performance in statistics courses has been extensively documented. Gal, Ginsburg, and Schau (1997) noted that, "students' attitudes and beliefs can impede (or assist) learning statistics, and may affect the extent to which students will develop useful statistical thinking skills and apply what they have learned outside the classroom." Statistics anxiety has previously been found to be the best predictor of achievement in statistics courses (Fitzgerald et al., 1996) and exhibits a consistently negative relationship with course performance (Onwuegbuzie et al. 2000; Benson 1989). Ruggeri et al. (2008) stated that "large numbers of students do experience statistics anxiety (and) have negative attitudes toward statistics", noting that more work is needed to understand the problems in order to address them.

Additionally, positive attitudes towards statistics appear to contribute to success in statistics courses (Roberts & Saxe 1982; Wise 1985) and enable students "to develop statistical thinking skills, to apply knowledge acquired in everyday life, and to have an enjoyable experience throughout the course." (Judi et al, 2011)

There have been a number of measurement instruments introduced to examine student's attitudes towards statistics, the first of which was the Statistics Attitudes Survey (SAS, Roberts & Bilderback 1980), a unidimensional scale with 33 items. The SAS has drawn criticism from numerous quarters for both its one-dimensional nature, which is not a viable assumption according to many attitude theories (Albarracin, Johnson, & Zanna, 2005), as well as its items, some of which test students' statistics knowledge instead of attitudes (Wise, 1985).

The Attitudes Towards Statistics scale (ATS, Wise 1985) is a two-dimensional survey that included a *Field* scale (20 items), measuring a students' beliefs about the use of statistics in their field of study; as well as a *Course* scale (9 items) that examined attitudes towards statistics in the course in which they were enrolled.

The original 28-item Survey of Attitudes Towards Statistics (SATS-28, Schau et al. 1995) was designed to measure four different aspects of students' attitudes towards statistics. The first of these aspects (*Affect*) measures students' feelings concerning statistics; *Cognitive Competence* examines students' attitudes about their intellectual knowledge and skills when applied to statistics; The *Value* dimension measures students' attitudes about the usefulness, relevance, and worth of statistics in their life; and finally attitudes about the difficulty of statistics as a subject is measured on the *Difficulty* dimension. More recently (Schau, 2003), a 36-item SATS-36 scale has also been introduced, which adds components measuring students' level of interest in statistics (*Interest*) and the amount of work the student dedicates to the learning of statistics (*Effort*).

An additional scale measuring statistics anxiety is the Statistics Anxiety Rating Scale (STARS, Cruise et al. 1985), a 51-item questionnaire measuring six different dimensions of statistics anxiety: (1) *Worth of Statistics*, measuring how useful students perceive statistics to be, (2) *Interpretation Anxiety* measuring the level of anxiety a student feels when interpreting statistical data or results; (3) *Test and Class Anxiety*; (4) *Computation Self-concept* measures a students' perception of their own math ability and anxiety about working math problems; (5) *Fear of Asking for Help* measures a student's anxiety over asking for help with items such as, "asking one of your professors for help in understanding a printout"; (6) *Fear of Statistics Teachers* measures a student's perception of statistics teachers with items such as, "statistics teachers talk so fast you cannot logically follow them".

Numerous studies have determined a relationship between statistics assessment outcomes and statistics attitudes, regardless of the measurement instrument used (Wise, 1985; Cashin & Elmore, 2005; Roberts & Reece, 1987; Dempster & McCorry, 2009). Mji and Onwuegbuzie (2004) correlated STARS with the ATS, finding that "after applying the Bonferroni adjustment (α=.05/6=.0083), the Course subscale scores were statistically significant and negatively correlated with five of the six anxiety dimensions." Additionally, Roberts and Reece (1987) found a correlation between the SAS and the ATS of .88, whereas Schau et al. (1995) reported correlations between the SATS-28 and the two ATS subscales Field and Course, respectively, of .34 and .79 on the Affect scale; .38 and .76 with Cognitive Competence; .76 and .40 with the Value dimension; and -.03 and .42 with the SATS Difficulty scale.

Emmioğlu and Capa-Aydin (2012) also reported coefficient alpha values ranging (depending on the sample used) from .80 to .85 for Affect (six items), from .77 to .82 for Cognitive Competence (six items), from .78 to .90 for Value (nine items), from .64 to .75 for Difficulty (seven items), from .80 to .84 for Interest (four items), and from .76 to .81 for Effort (four items).

CHAPTER 3

METHODS

Instruments Used

For this study, I wanted to collect information on students' motivations for attending college, their attitude towards statistics, as well as their conceptual understanding of statistics at the end of an introductory statistics course. I also intended to collect some demographic information, such as gender, from each participating student.

To measure motivation for attending college, I used the 28-item AMS (Vallerand et al., 1992), which includes seven dimensions – three aspects of intrinsic motivation, three aspects of extrinsic motivation, as well as amotivation. These 28 items, along with their seven respective subscales, are shown in Appendix D. Attitude towards statistics was measured by the 28-item SATS-28 (Schau et al., 1995). This shorter, four-dimension version of SATS was chosen (rather than the longer, six-dimension SATS-36) due to it already being administered in the STAT 2000 introductory statistics course at the University of Georgia, where I collected my data. These 28 questions, along with their four respective sub-scales, are shown in Appendix B.

The demographic variables on which I collected information were gender, whether the course was required for the student, the age of the student, year of college (freshman/ sophomore/ junior/ senior/ other), and the number of college mathematics and/or statistics courses taken by the student, prior to the current semester. These five questions, which were appended to the SATS survey form, are shown in Appendix C.

Finally, conceptual understanding of statistics was measured by the student's score on the 40-item CAOS 4 test. For copyright reasons, the actual questions used on the CAOS4 test can't be published in this dissertation, but the statistical concepts examined in each of the 40 questions are shown in Appendix A.

To motivate complete answers to all instruments, the AMS, SATS and demographic information were assigned as part of two separate lab assignments during the first few weeks of the Fall 2012 semester. The full, 40-item CAOS 4 test was given during the penultimate week of the semester, with each correct answer rewarded with a 0.05 increase in the student's course average (out of 100), meaning that a perfect score on the test would result in a two-percentage-point increase in the student's course grade. The intent of this reward system was to motivate honest, considered responses to all CAOS questions and survey items.

The STAT 2000 course at UGA is a 4 credit hour course, comprised of 3 credit hours of lecture and a separate credit hour in the form of a weekly, 50-minute computer lab. All survey instruments were completed in the lab, as part of the lab exercise for that week, whereas the CAOS test was given instead of the regular exercise for that week, also in the lab.

A list of the survey instruments that were considered is included in Table 1, with further clarification of the instruments used provided in Table 2.

Research Questions

For the first part of this study, I planned to test the suitability of the alternative configuration of the AMS specified by Smith et al. (2012) and to compare it with the full configuration through the use of principal component analysis. Whilst various studies have seemed to confirm the appropriateness of the original seven-factor model, others have suggested a five-factor model (with only one intrinsic motivation factor), whilst Smith et al. (2012) suggested an alternative four-factor configuration with one overall intrinsic motivation factor, one amotivation factor and two extrinsic motivation factors (of external and identified regulation). The later configuration does not include all items in the original AMS.

After determining a suitable configuration, I then explored the effect of motivation and attitude towards statistics on conceptual understanding. A structural equation model was constructed to describe the effects of the (potentially seven) motivation and (four) attitude subscales on conceptual understanding (measured by total score on the CAOS test). A multiple regression model was also constructed to look at the effect these constructs have on conceptual understanding.

The final part of my analysis looks at whether certain types of statistics misconceptions were more common for students with differing types of motivation. An item analysis was performed looking at each item individually and examining whether there were any significant performance differences between students with, for example, high and low external regulation.

Data Collection

Data were collected from 1350 students enrolled in the STAT 2000 introductory statistics course at the University of Georgia in the Fall 2012 semester. In the second week of the course, students were asked to complete the 28-item SATS survey, in addition to several demographic questions. In week 3, students then completed the AMS. Finally, in the penultimate week of the course, the students completed the CAOS test.

Any students with three or more missing answers on the SATS or AMS were removed from the analysis. If a student's response form included only one or two missing values, the average value for that item over all respondents was used to replace the missing value. Whereas this 'mean substitution' method is typically frowned upon in statistics literature, the presence of multiple items measuring each subscale should still allow a reasonably accurate measure on each subscale to be formed for any participant, so long as there are not more than a couple of missing values for that participant. Missing values on the CAOS 4 test of conceptual understanding were left as missing values and graded as incorrect, since the absence of a response on a test question can often be taken to mean the student did not know the answer.

The data included 1263 completions of the SATS survey, 1282 complete AMS responses, and 1202 completions of the CAOS test. A total of 1094 students completed all survey instruments as well as the CAOS test. A breakdown of the number of students completing each instrument and the number of responses that were removed is included in Table 3.

The students in the sample were predominantly female (66%) and between the ages of 18-22 (95%). When asked about college mathematics and/or statistics courses the students had taken prior to the current semester, 34% had not taken any, the same number had taken one, and a further 23% had taken two. Additionally, 25% of the students were freshmen, 40% sophomores, 24% juniors and 9% seniors, with 2% answering 'other'. The demographic breakdown of the sample is shown in Table 4.

With the motivation for attending college subscales, no reverse coding was necessary for any item, since all items were measured on a Likert 1-7 scale, with a value of seven being indicative of a greater presence of the construct being measured. With the statistics attitude subscales, values were again measured on a Likert seven-point scale, but negatively worded items were reverse coded (a one coded as a seven, a two as a six, etc.) so larger values were more indicative of a positive attitude on one of the measured constructs.

Summary statistics for the AMS data collected from the sample is provided in Table 5; with Table 6 showing similar information for the SATS-28 items, after the reverse coding had taken place. Percentages correct for each item on the CAOS 4 test (Table 7) and summary statistics for the overall CAOS 4 scores (Table 8) are also provided.

Table 1

Summary of Survey Instruments Considered

Conceptual Understanding	Motivation	Statistics Attitudes
SCI (2003)	MSLQ (1990)	SAS (1980)
SRA (1998)	**AMS (1992)**	ATS (1985)
CAOS 4 (2005)		**SATS-28 (1995)**
		SATS-36 (2003)
		STARS (1985)

Survey instruments selected for use in the study are shown in bold.

Table 2

Summary of Survey Instruments Used

Instrument	Items	Summary of Instrument	Appendix
Academic Motivation Survey (AMS)	28	7* subscales measuring: Intrinsic Motivation to Know (4 items) Intrinsic Motivation to Accomplish Things (4 items) Intrinsic Motivation to Experience Stimulation (4 items) External Regulation (4 items) Introjected Regulation (4 items) Identified Regulation (4 items) Amotivation (4 items)	D
Survey of Attitudes towards Statistics (SATS-28)	28	4 subscales measuring: Affect (6 items) Cognitive Competence (6 items) Value (9 items) Difficulty (7 items)	B
Demographic Questions	5	Separate questions measuring (1) gender, (2) whether the course was required, (3) age, (4) year of college (freshman/ sophomore/ junior/ senior/ other), (5) number of college mathematics and/or statistics courses taken by the student, prior to the current semester.	C
The Comprehensive Assessment of Outcomes in Statistics (CAOS 4)	40	Separate questions measuring conceptual understanding at the end of an introductory statistics course	A

* Vallerand et al. (1992) specified 7 subscales, but numerous studies have raised doubts about the true dimensionality.

Table 3

Summary of Survey Instrument Completion

Instrument	Number of Responses
Academic Motivation Survey (AMS)	1284 students attempted survey instrument 2 students removed for 3-or-more missing values 4 students left one missing value, which was estimated by imputation 1282 responses used in analysis of AMS
Statistics Attitude Survey (SATS-28)	1263 students attempted survey instrument 0 students removed for 3-or-more missing values 2 students left one missing values, which was estimated by imputation 1263 responses used in analysis of SATS-28
Demographic Questions	1263 students attempted survey instrument 0 students removed for 3-or-more missing values 1 student left one missing value, which was estimated by imputation 1263 responses used in analysis of demographic data
The Comprehensive Assessment of Outcomes in Statistics (CAOS)	1202* students attempted CAOS test 88 students left at least one missing value which were included in the analysis as an incorrect answer – of the 88, 71 left a single missing answer, 12 left two blanks, with the other five students leaving 3, 4, 9, 12 and 12 blanks out of 40 questions. 1202 responses used in analysis of CAOS test
All instruments	1350 students attempted at least one survey instrument 68 students removed for incomplete AMS 70 additional students removed for incomplete SATS-28 0* additional students removed for incomplete demographics 118** additional students removed for non-attempt at CAOS test 1094 responses complete across all instruments

* This is not a surprise, since the demographic questions were asked at the end of the SATS-28 survey.

** Since the CAOS test was assigned towards the end of the course, the high number of CAOS non-attempts from students that had previously completed all survey instruments was presumably due to course withdrawals.

Table 4

Demographic Breakdown of Sampled Students

Question	Responses	Frequency
Q1. What is your gender?	Male	425 (33.7%)
	Female	838 (66.3%)
Q2. What is your year of college?	Freshman	318 (25.2%)
	Sophomore	501 (39.7%)
	Junior	303 (24.0%)
	Senior	118 (9.3%)
	Other	23 (1.8%)
Q3. Is this course required for your college degree?	Yes	1109 (87.8%)
	No	154 (12.2%)
Q4. Which of these age groups do you belong to?	17 or under	15 (1.2%)
	18-22	1193 (94.5%)
	23-28	42 (3.3%)
	29-34	6 (0.005%)
	35+	6 (0.005%)
Q5. How many college mathematics and/or statistics courses have you completed, not counting this semester?	0	427 (33.8%)
	1	428 (33.9%)
	2	287 (22.7%)
	3	78 (6.2%)
	4+	43 (3.4%)

A total of N = 1263 students completed the demographic survey. One student failed to answer question 4.

Table 5

Summary Statistics for Academic Motivation Scale (AMS) Items

Variable	*N*	*M*	*SD*
Intrinsic Motivation To Know			
IMTK1	1282	5.259	1.269
IMTK2	1282	5.105	1.435
IMTK3	1282	5.399	1.370
IMTK4	1282	5.528	1.321
Intrinsic Motivation To Accomplish Things			
IMTA1	1282	4.590	1.490
IMTA2	1282	4.936	1.507
IMTA3	1281	4.776	1.542
IMTA4	1282	4.937	1.534
Intrinsic Motivation To Experience Stimulation			
IMTS1	1281	4.137	1.548
IMTS2	1282	3.975	1.766
IMTS3	1281	3.703	1.751
IMTS4	1282	3.821	1.756
External Regulation			
ER1	1282	5.525	1.495
ER2	1282	6.041	1.271
ER3	1282	5.610	1.519
ER4	1282	5.803	1.348
Introjected Regulation			
IR1	1282	5.191	1.663
IR2	1281	5.039	1.595
IR3	1282	5.028	1.593
IR4	1282	5.291	1.507
Identified Regulation			
IDENT1	1282	6.252	1.204
IDENT2	1282	6.095	1.207
IDENT3	1282	5.516	1.381
IDENT4	1282	5.590	1.390
Amotivation			
AMOT1	1282	1.945	1.402
AMOT2	1282	1.979	1.486
AMOT3	1282	1.675	1.317
AMOT4	1282	1.803	1.367

All items measured on a 1 to 7 scale, with larger values indicating the item corresponds with the student’s reason for attending college.

Table 6

Summary Statistics for Student Attitudes towards Statistics (SATS-28) Items

Variable	*N*	*M*	*SD*
Affect			
AFFECT1	1263	4.352	1.343
AFFECT2 *	1263	4.389	1.631
AFFECT3 *	1263	4.549	1.655
AFFECT4 *	1263	4.050	1.651
AFFECT5	1263	4.008	1.424
AFFECT6 *	1263	4.523	1.881
Cognitive Competence			
COGCOMP1 *	1263	4.627	1.594
COGCOMP2 *	1263	5.403	1.491
COGCOMP3 *	1263	4.096	1.566
COGCOMP4	1263	5.926	1.223
COGCOMP5	1263	5.246	1.299
COGCOMP6 *	1263	4.553	1.487
Value			
VALUE1 *	1263	5.910	1.269
VALUE2	1263	4.308	1.608
VALUE3	1263	5.016	1.390
VALUE4 *	1263	5.340	1.328
VALUES5 *	1263	5.243	1.424
VALUE6	1263	4.311	1.526
VALUE7 *	1263	5.338	1.362
VALUE8 *	1262	5.335	1.416
VALUE9 *	1263	5.367	1.401
Difficulty			
DIFFIC1	1263	4.405	1.353
DIFFIC2 *	1263	3.927	1.451
DIFFIC3	1263	3.600	1.248
DIFFIC4 *	1263	3.583	1.266
DIFFIC5 *	1263	4.516	1.422
DIFFIC6 *	1263	3.951	1.193
DIFFIC7 *	1263	3.973	1.268

All items measured on a 1 to 7 scale, with larger values indicating a positive attitude on the item

* Score shown is after the variable was reverse-coded

Table 7

Number and Percentage of Students Answering each CAOS 4 Item Correctly

Item	Number (and Percentage) Correct		Item	Number (and Percentage) Correct	
1	947	(78.8%)	21	985	(81.9%)
2	627	(52.2%)	22	571	(47.5%)
3	1009	(83.9%)	23	711	(59.2%)
4	799	(66.5%)	24	865	(72.0%)
5	955	(79.5%)	25	581	(48.3%)
6	210	(17.5%)	26	738	(61.4%)
7	53	(4.4%)	27	608	(50.6%)
8	835	(69.5%)	28	498	(41.4%)
9	203	(16.9%)	29	714	(59.4%)
10	242	(20.1%)	30	394	(32.8%)
11	1134	(94.3%)	31	1003	(83.4%)
12	1092	(90.8%)	32	112	(9.3%)
13	881	(73.3%)	33	349	(29.0%)
14	705	(58.7%)	34	726	(60.4%)
15	714	(59.4%)	35	582	(48.4%)
16	484	(40.3%)	36	659	(54.8%)
17	744	(61.9%)	37	162	(13.5%)
18	1006	(83.7%)	38	380	(31.6%)
19	863	(71.8%)	39	103	(8.6%)
20	1157	(96.3%)	40	731	(60.8%)

Percentages calculated based on the *N*=1202 students that attempted the CAOS test.

Table 8

Summary Statistics for Overall CAOS 4 Scores

N	1202	P1	13.000
M	21.740	P5	16.000
Mdn	21.000	P10	17.000
Mode	21.000	P25	19.000
SD	4.080	P75	24.000
Min	9.000	P90	27.000
Max	35.000	P95	29.000
Range	26.000	P99	33.000

CHAPTER 4

ANALYSIS

Analysis 1. Principal Component Analysis to Determine Dimensionality of Academic Motivation Scale (AMS)

For the first part of my analysis I wanted to determine the most appropriate dimensionality of the Academic Motivation Scale. A total of 1284 students took the survey, however two respondents were removed completely based on having three-or-more missing values. Four additional surveys contained a single missing item, for which the missing response was changed to the average response for that item.

I conducted a principal component analysis using a Varimax rotation procedure to identify the appropriate dimensionality of the 28 items. This procedure was performed in SAS Version 9.3. I set a minimum eigenvalue criteria of one, and for an item to be included as part of a scale it needed to have a loadings of .5 or higher on that factor, and not load higher than .3 on any other factor, to ensure reliability of the factor subscales. These cutoffs were the same ones used by Smith et al. (2012) and were chosen to see if that solution would be replicated.

The resulting model consisted of four factors that corresponded very closely with the model identified by Smith et al. (2012), with factors of Intrinsic Motivation, Amotivation, External Regulation and Identified Regulation. The External Regulation factor initially specified by my model consisted of three (rather than six) items, due to the other three items having slightly higher cross-loadings on other factors. Additionally, my Identified Regulation factor has three items (rather than two) due to the lack of a significant cross-loading for the IR4 item. The Intrinsic Motivation and Amotivation factors correspond with the findings of Smith et al. (2012) exactly.

Examination of the alpha coefficients revealed a problem with my External Regulation dimension, with an alpha value of .768 and item-deleted alpha values ranging from .622 to .790. To remedy this problem, one of the External Regulation factors present in the Smith et al. (2012) model but deleted from mine (ER3) was added back in, since this was the omitted item with the strongest correlation with the factor and was only initially removed due to a cross-loading of .31 with Identified Regulation.

It is notable that all four of the items measuring identified regulation were excluded from the model due to significant cross-loadings across several factors. Additionally, all four of the items measuring intrinsic motivation to accomplish were excluded due to significant cross-loadings on both the intrinsic motivation and identified regulation factors.

The final results of this factor analysis are included in Table 9 with further summary factor information and alpha values in Table 10.

Table 9

Factor Loadings for Academic Motivation Scale Items

Factor and Item	Factor Loadings 1	2	3	4
Factor 1: Intrinsic Motivation				
IMTS2 For the pleasure I experience when I read interesting authors	**.81**	.13	.04	.02
IMTS3 For the pleasure I experience when I feel completely absorbed by what certain authors have written	**.80**	.20	.02	.07
IMTS4 For the ‘high’ feeling that I experience while reading about interesting subjects	**.73**	.14	-.03	.24
IMTK2 For the pleasure I experience when I discover new things never seen before	**.72**	-.16	.09	.21
IMTS1 For the intense feelings I experience when I am communicating my own ideas to others	**.70**	.06	.07	.13
IMTK1 Because I experience pleasure and satisfaction while learning new things	**.64**	-.29	.12	.07
Factor 2: Amotivation				
AMOT1 Honestly, I don’t know; I really feel that I am wasting my time in school	-.08	**.83**	-.14	-.10
AMOT2 I once had good reasons for going to college; however, now I wonder whether I should continue	.00	**.82**	-.11	-.07
AMOT3 I can’t see why I go to college and frankly, I couldn’t care less	.02	**.82**	-.13	-.08
AMOT4 I don’t know; I can’t understand what I am doing in school	-.05	**.82**	-.16	-.09
Factor 3: External Regulation				
ER4 In order to have a better salary later on	-.01	-.12	**.80**	.24
ER2 In order to obtain a more prestigious job later on	.06	-.23	**.79**	.18
ER3 I want to have ‘the good life’ later on	.06	-.04	**.73**	**.31**
ER1 Because with only a high-school degree I would not find a high paying job later on	.02	.00	**.69**	.01
Factor 4: Identified Regulation				
IR3 To show myself that I am an intelligent person	.21	-.06	.24	**.79**
* IR4 Because I want to show myself that I can succeed in my studies	.28	-.16	.17	**.79**
IR1 To prove to myself that I am capable of completing my college degree	.25	-.07	.20	**.70**

Excluded Items (***due to cross-loadings highlighted in bold italics***)				
IMTK3 For the pleasure that I experience in broadening my knowledge about subjects that appeal to me	***.68***	***-.30***	.12	.21
IMTK4 Because my studies allow me to continue to learn about many things that interest me	***.60***	***-.33***	.10	***.31***
IMTA1 For the pleasure I experience while surpassing myself in my studies	***.62***	-.13	.10	***.31***
IMTA2 For the pleasure that I experience while I am surpassing myself in one of my personal accomplishments	***.61***	-.14	.13	***.46***
IMTA3 For the satisfaction I feel when I am in the process of accomplishing difficult academic activities	***.60***	-.14	.09	***.47***
IMTA4 Because college allows me to experience a personal satisfaction in my quest for excellence in my studies	***.58***	-.16	.16	***.53***
IR2 Because of the fact that when I succeed in college I feel important	***.30***	-.04	***.35***	***.64***
IDENT1 Because I think that a college education will help me prepare for a career that I have chosen	.20	***-.52***	***.49***	.12
** IDENT2 Because eventually it will enable me to enter the job market in a field that I like	.12	***-.45***	***.62***	.07
** IDENT3 Because this will help me make a better choice regarding my career orientation	***.31***	-.27	***.54***	.29
IDENT4 Because I believe that a few additional years of education will improve my competence as a worker	.25	-.25	***.42***	***.40***

* Excluded due to cross-loading in Smith et al. (2012) model

** Classified as part of External Regulation factor in Smith et al. (2012) model

Table 10

Eigenvalues and Percentage of Variance Explained by Each Factor, along with Summary Statistics and Consistency Measures for Included Items

Factor/ Variable	Eigenvalue	% of Total Variance Explained	*M*	*SD*	Item-Total Correlation	Alpha Coefficient Item Deleted
Intrinsic Motivation (α = .860)	6.043	21.583				
IMTS2			3.975	1.766	.741	.819
IMTS3			3.703	1.751	.722	.823
IMTS4			3.821	1.756	.684	.830
IMTK2			5.105	1.435	.618	.842
IMTS1			4.137	1.548	.622	.841
IMTK1			5.259	1.269	.519	.859
Amotivation (α = .886)	3.886	13.879				
AMOT1			1.945	1.402	.763	.849
AMOT2			1.979	1.486	.738	.858
AMOT3			1.675	1.317	.745	.856
AMOT4			1.803	1.367	.757	.851
External Regulation (α = .820)	3.799	13.569				
ER4			5.803	1.348	.734	.730
ER2			6.041	1.271	.679	.756
ER3			5.610	1.519	.655	.768
ER1			5.525	1.495	.509	.833
Identified Regulation (α = .838)	3.716	13.271				
IR3			5.028	1.593	.703	.773
IR4			5.291	1.507	.742	.735
IR1			5.191	1.663	.659	.816

Analysis 2. Construction of SEM Model to Explain Conceptual Understanding of Statistics

For the second part of my analysis, I constructed a structural equations model to look at the relationship between motivation for attending college, statistics attitudes and conceptual understanding of statistics. Motivation for attending college was measured using the four AMS dimensions determined earlier of Intrinsic Motivation, Amotivation, External Regulation, and Identified Regulation. To measure statistics attitudes, I used the four dimensions of Affect (measured by 6 items), Cognitive Competence (6 items), Value (9 items) and Difficulty (7 items) measured by the SATS-28. Conceptual understanding of statistics in this model was measured by the students' overall score on the CAOS 4 test out of 40.

My hypothesized model had the four motivation subscales correlated with each other, as discussed by Smith et al. (2012), and the four statistics attitude subscales correlated with each other (Schau et al., 1995) but none of the motivation subscales correlated with any of the attitude subscales. All attitude and motivation subscales were expected to correlate with conceptual understanding of statistics, as measured by score on the CAOS test.

To construct the model, I used Mplus Version 4.0. Since all of the motivation and statistics attitude scale items are ordinal and measured on a seven-point Likert scale, with many of the items exhibiting skewedness, a robust maximum likelihood estimation technique (MLR) was used. This is the appropriate estimation technique to use with non-normal ordinal data with multi-level responses (Hu, Bentler, & Kano, 1992), since the chi-square test statistic is usually inflated with non-normal samples, and the standard errors usually too small (Curran, West, & Finch, 1996). This scaling of the chi-square test statistic and standard errors does not affect the parameter estimates or unique variances. The resulting model is illustrated in Figure 1.

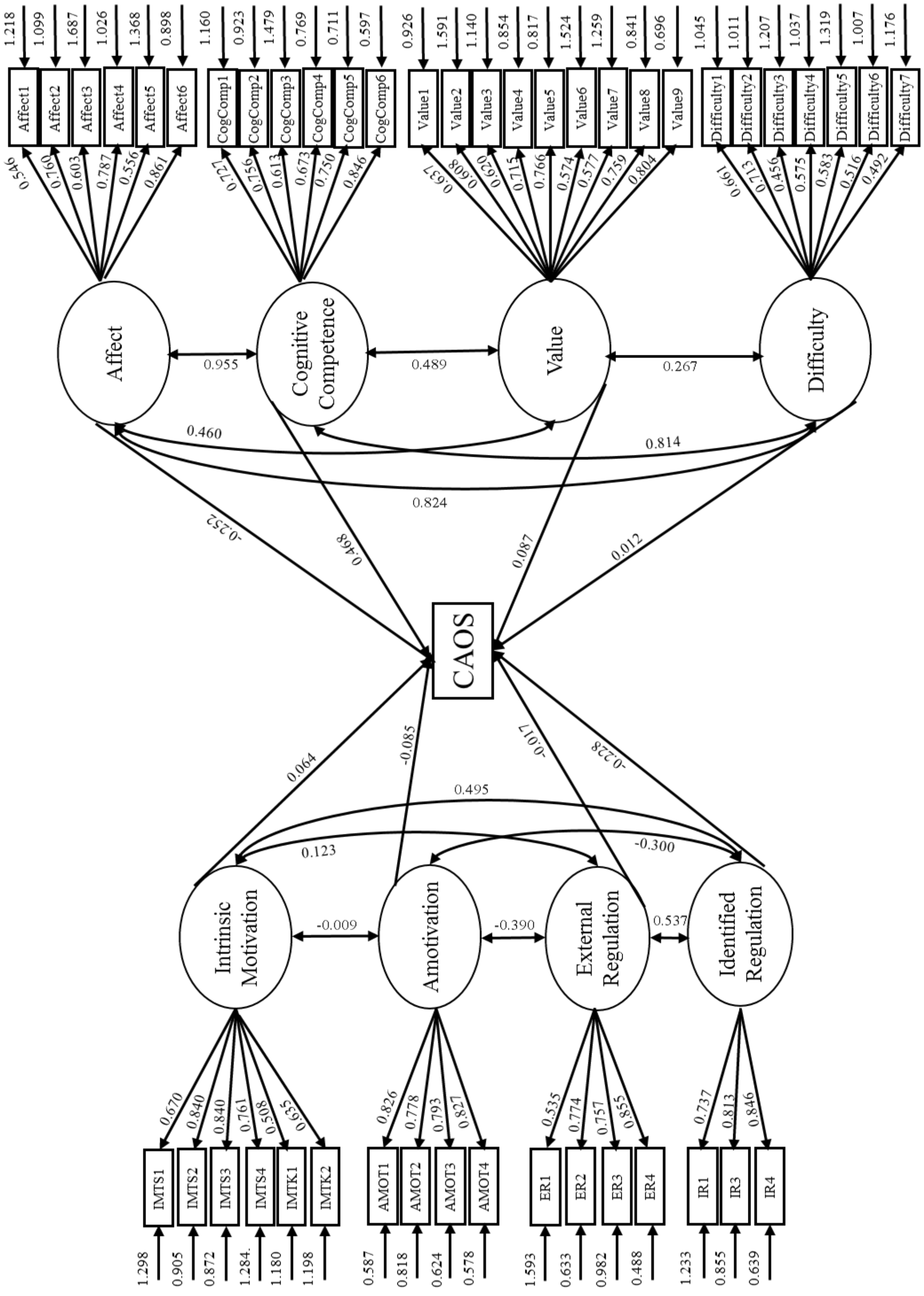

Figure 1

Structural equations model of college motivation and statistics attitude subscales and conceptual understanding of statistics.

This model has a chi-square value of 3466.007 (df = 970) which indicates a bad fit. However, this test is very sensitive to sample size and nearly always rejects the null hypothesis of a good fit when the sample size is large (Hooper et al., 2008). One suggested remedy to this problem is the relative/normed chi-square (Wheaton et al., 1977) which is calculated by dividing the chi-square statistic by the degrees of freedom to give a value for our model of 3.573. Although there is no universally accepted cutoff for acceptability of this statistic (Hooper et al., 2008), this ratio does fall within the cutoff of 5 originally recommended by Wheaton (1977).

Additionally, the root mean square error of approximation (*RMSEA*) for the model of .049 falls below the most stringent cutoff of .05 typically suggested in the literature (Hooper et al. 2008), further indicating a good fit to the data. The RMSEA measures the lack of fit of the model to the population data and is calculated as:

$$\text{RMSEA}=\sqrt{\frac{\hat{\delta}_{Model}}{df_{Model}(n-1)}}$$

where δ measures the degree of misspecification of the model and is estimated as $\hat{\delta}_i = \chi_i^2 - df_i$.

The Comparative Fit Index (CFI) compares the performance of a model to the null model that assumes zero correlation between all observed variables. It is calculated as:

$$\text{CFI} = 1 - \frac{\hat{\delta}_{Model}}{\hat{\delta}_{Null}}$$

The model's *CFI* of .875 does raise questions about possible misspecification, falling below both the recommended cutoffs of .9 (used in earlier studies) and .95 (the more recent standard) (Hu & Bentler, 1999). The Mplus code used to generate this model is included in Appendix E.

To try and improve fit, a number of other models were considered. When paths were added to correlate the motivation factors with the statistics attitude factors, there was a slight improvement in the fit measures (normed χ^2 = 3.54, *RMSEA* = .048, *CFI* = .879) however these paths did not seem intuitive and the improvement in fit by adding these extra paths seemed negligible.

Additionally, removing items from subscales with which they shared a low item-total correlation offered some improvement, but not enough to warrant the reduction in scope of the subscale. Removing ER1 from the External Regulation subscale produced a normed chi-square of 3.65, *RMSEA* of .049 and a *CFI* of .875. Removing the two intrinsic motivation to experience stimulation (IMTS1 & IMTS2) from the intrinsic motivation subscale produced a normed chi-square of 3.47, *RMSEA* of .048 and a *CFI* of .882. Both of these adjustments together lead to a normed chi-square of 3.54, *RMSEA* of .048 and a *CFI* of .882.

The most notable attempt at improving fit though came with the consideration of the original seven-subscale configuration of the AMS in the model. This model produced a normed chi-square of 2.99, *RMSEA* of .043 and a *CFI* of .890. The addition of paths between the motivation and statistics attitude subscales with this model lead to a normed chi-square of 2.97, *RMSEA* of .042 and a *CFI* of .893. It should be noted however that these slight improvements to fit came at the expense of the splitting of motivation into seven subscales (a solution deemed unnecessary earlier) and through the addition of additional, unintuitive paths, and as a result I see no reason to move away from the originally specified model.

Analysis 3. Construction of Multiple Regression Models to Explain Conceptual Understanding of Statistics

For the third part of my analysis, I constructed a multiple regression model, utilizing the Stepwise model building procedure in SAS Version 9.3, to predict conceptual understanding of statistics, as measured by score on the CAOS 4 test. Only variables significant at an $\alpha = .15$ level of significance were considered for entry into the model (*SLE* = .15) and to remain in the model, the variable needed to be significant at the $\alpha = .10$ level (*SLS* = .10). The initial variables I considered for entry into the model were the average scores on the four subscales measuring motivation for attending college (Intrinsic Motivation, Amotivation, External Regulation, and Identified Regulation), and the average scores on the four statistics attitude subscales (Affect, Cognitive Competence, Value, and Difficulty).

The method of averaging the scores across all items measuring a construct was selected due to each factor being measured by a different number of items. This would produce a total of eight different potential explanatory variables, each measured on a 1-7 scale. This averaging also took place after the negatively worded items on the SATS-28 had been reverse coded.

Subsequent models under consideration will also include the five demographic variables discussed in Table 4, coded as follows:

Female – Coded 1 if the student was Female, 0 if Male.

Class – Coded 0 if the student was a Freshman, 1 if a Sophomore, 2 if a Junior, 3 if a Senior. Since the questionnaire had a fifth choice of ‘other’, these responses were left as blanks in the analyzed data due to it being unclear where these students would fall on the continuum.

Required – Coded 1 if the course was a required course for the student, 0 otherwise.

OlderStudent – Coded 0 if student was 22-or-under, 1 if student 23-or-older. Note that the five initial age group categories were split into groups of either '22-or-under' or '23-or-older' due to there being so few responses in many of the older categories.

Coll_Math&Stat – The number of college mathematics and/or statistics courses the student had completed, not counting the current semester. Completion of 4-or-more courses was coded just as 4.

Summary statistics for these eight potential explanatory variables and five demographic variables, along with their respective correlation with the overall CAOS 4 score, are shown in Table 11. A full table of correlation coefficients is provided in Appendix F.

Table 11

Summary Statistics for AMS and SATS-28 Subscales and Demographic Variables

Variable	*M*	*SD*	Min	Max	CAOS 4 *r*
Motivation					
Intrinsic Motivation	4.342	1.219	1.000	7.000	-0.029
Amotivation	1.804	1.168	1.000	7.000	-0.068
External Regulation	5.757	1.125	1.000	7.000	-0.072
Identified Regulation	5.179	1.369	1.000	7.000	-0.164
Statistics Attitude					
Affect	4.349	1.194	1.000	7.000	0.230
Cognitive Competence	5.008	1.101	1.000	7.000	0.265
Value	5.135	1.004	1.000	7.000	0.187
Difficulty	4.009	0.854	1.143	7.000	0.195
Demographic					
Female	0.674	0.469	0.000	1.000	-0.213
Class*	1.110	0.894	0.000	3.000	-0.049
Required	0.872	0.334	0.000	1.000	-0.117
Older Student	0.032	0.176	0.000	1.000	0.001
College Math & Stat	1.093	1.047	0.000	4.000	-0.092

Since these item averages will be used in a multiple regression predicting CAOS score, they are only calculated based on the n = 1094 students that completed all instruments.

* Statistics for Class variable calculated based on 1075 responses after 'other' responses were turned to blanks.

The initial model chosen by the Stepwise procedure, which only considered the four motivation averages and four statistics attitude averages for inclusion, was highly significant ($F(5, 1088) = 27.46$, $p < .0001$) and contained three motivation factors (Identified Regulation, Amotivation, and Intrinsic Motivation) and two statistics attitude factors (Cognitive Competence and Value). The model had a RMSE of 3.850. It may be initially surprising that the attitudinal factor of Affect is not present in the model, however since the correlation between Affect and Cognitive Competence is .828 (Appendix F), Affect was never entered into the model since Cognitive Competence was entered first.

A table of parameter estimates is shown in Table 12 as well as a summary of the Stepwise selection process in Table 13.

Table 12

Table of Parameter Estimates for Stepwise Regression Model Predicting CAOS Score from Motivation and Statistics Attitudes Subscales.

Variable	b_i^*	*SE*	Type II SS	*F*	*p*
Intercept	18.54375	0.92621	5944.51212	400.84	< .0001
Intrinsic	0.19651	0.10908	48.13474	3.25	.0719
Amotivation	-0.21555	0.10608	61.22765	4.13	.0424
IdentiReg	-0.63683	0.09983	603.50890	40.70	< .0001
CogComp	0.78034	0.11767	652.19525	43.98	< .0001
Value	0.40392	0.13075	141.52098	9.54	.0021

Table 13

Summary of Stepwise Selection Process for Predicting CAOS Score from Motivation and Statistics Attitudes Subscales

Step	Variable Entered	Variable Removed	Number Vars In	Partial R^2	Model R^2	*C(p)*	*F*	*p*
1	CogComp		1	.0704	.0704	47.350	82.70	< .0001
2	IdentiReg		2	.0257	.0961	17.866	31.06	< .0001
3	Value		3	.0104	.1065	7.129	12.70	.0004
4	Amotivation		4	.0029	.1094	5.626	3.50	.0616
5	Intrinsic		5	.0026	.1121	4.385	3.25	.0719

For the next part of my multiple regression analysis, I looked into the inclusion of potential interactions between subscales. A similar Stepwise selection procedure was conducted, this time with interactions between all subscales considered for inclusion, alongside all individual subscales. The resulting model was again highly significant ($F(7, 1086) = 23.04$, $p < .0001$) with the R^2 for the final model rising slightly to .1293 from the .1121 of the previous, non-interaction model. The RMSE also dropped from 3.850 to 3.817.

A table of parameter estimates (Table 14) and a summary of the stepwise selection process (Table 15) are once again included. Also, since the statistics literature is clear that interaction effects should never be included without the main effects also being present in the model, further attempts were made to refine the model by including all main effects that are present in any significant interactions, even if the main effect is not significant. Any interactions that subsequently became non-significant (in this case the interaction between Identified Regulation and Amotivation with $p = .4650$) were dropped from the final model, shown in Table 16. This model has an R^2 value of .1296, a RMSE of 3.822 and is also highly significant ($F(10, 1083) = 16.13$, $p < .0001$).

Table 14

Table of Parameter Estimates for Stepwise Regression Model Predicting CAOS Score from Motivation and Statistics Attitudes Subscales and Interactions

Variable	b_i^*	*SE*	Type II SS	*F*	*p*
Intercept	18.87127	0.93698	5909.71111	405.64	< .0001
IdentiReg	0.86141	0.32386	103.06660	7.07	.0079
Value	-0.47338	0.26276	47.28654	3.25	.0719
IdentiReg*Intrinsic	-0.17063	0.04943	173.57384	11.91	.0006
Intrinsic*Value	0.20549	0.05089	237.49850	16.30	< .0001
Amotivation*IdentiReg	-0.04707	0.02100	73.21124	5.03	.0252
IdentiReg*Difficulty	-0.17022	0.05034	166.61756	11.44	.0007
CogComp*Difficulty	0.19769	0.03182	562.45228	38.61	< .0001

Table 15

Summary of Stepwise Selection Process for Predicting CAOS Score from Motivation and Statistics Attitudes Subscales and Interactions

Step	Variable Entered	Variable Removed	Number Vars In	Partial R^2	Model R^2	*C(p)*	*F*	*p*
1	CogComp* Value		1	.0721	.0721	69.518	84.84	< .0001
2	IdentiReg		2	.0312	.1033	32.484	38.01	< .0001
3	CogComp* Difficulty		3	.0040	.1073	29.498	4.87	.0275
4	Amotivation *IdentiReg		4	.0034	.1107	27.259	4.15	.0418
5	Intrinsic* Value		5	.0040	.1147	24.244	4.93	.0266
6	IdentiReg* Difficulty		6	.0045	.1192	20.627	5.55	.0187
7		IdentiReg	5	.0000	.1192	18.667	0.04	.8413
8		CogComp* Value	4	.0014	.1178	18.444	1.76	.1854
9	IdentiReg* Intrinsic		5	.0058	.1236	13.209	7.19	.0075
10	IdentiReg		6	.0031	.1267	11.290	3.90	.0484
11	Value		7	.0026	.1293	10.038	3.25	.0719
12	ExternalReg *IdentiReg		8	.0017	.1310	9.953	2.08	.1492
13		ExternalReg *IdentiReg	7	.0017	.1293	10.038	2.08	.1492

Table 16

Table of Parameter Estimates for Final Regression Model Predicting CAOS Score from Motivation and Statistics Attitudes Subscales and Interactions

Variable	b_i^*	*SE*	*t*	*p*
Intercept	16.01894	3.93525	4.07	< .0001
IdentiReg	1.13620	0.50069	2.27	.0234
Intrinsic	0.25454	0.56406	0.45	.6519
Amotivation	-0.22474	0.10535	-2.13	.0331
CogComp	0.07560	0.39324	0.19	.8476
Value	-0.34896	0.40252	-0.87	.3862
Difficulty	0.61886	0.78070	0.79	.4281
IdentiReg*Intrinsic	-0.18861	0.05943	-3.17	.0015
IdentiReg*Difficulty	-0.24043	0.09366	-2.57	.0104
Intrinsic*Value	0.17720	0.08774	2.02	.0437
CogComp*Difficulty	0.16453	0.09806	1.68	.0937

I next ran a Stepwise regression procedure that considered the motivation and statistics attitude averages, without interactions, alongside the demographic characteristics. The resulting model was once again highly significant ($F(7, 1086) = 26.96$, $p < .0001$) with the R^2 rising to .1481 from the .1121 of the motivation and attitude main effects model. The RMSE also dropped from 3.850 to 3.817.

A table of parameter estimates (Table 17) and a summary of the stepwise selection process (Table 18) are once again included.

Table 17

Table of Parameter Estimates for Stepwise Regression Model Predicting CAOS Score from Motivation and Statistics Attitudes Subscales and Demographic Characteristics

Variable	b_i^*	*SE*	Type II SS	*F*	*p*
Intercept	21.51779	1.01047	6457.76982	453.47	< .0001
IdentiReg	-0.4770	0.08753	422.96793	29.70	< .0001
Amotivation	-0.27088	0.10590	93.18126	6.54	.0107
CogComp	0.65774	0.11819	441.03505	30.97	< .0001
Value	0.35430	0.12959	106.44113	7.47	.0064
Female	-1.40247	0.25994	414.53624	29.11	< .0001
Required	-0.68262	0.35081	53.91989	3.79	.0519
Coll_Math&Stats	-0.40658	0.11225	186.84665	13.12	.0003

Table 18

Summary of Stepwise Selection Process for Predicting CAOS Score from Motivation and Statistics Attitudes Subscales and Demographic Characteristics

Step	Variable Entered	Variable Removed	Number Vars In	Partial R^2	Model R^2	*C(p)*	*F*	*p*
1	CogComp		1	.0689	.0689	87.913	79.46	< .0001
2	Female		2	.0275	.0964	55.721	32.59	< .0001
3	IdentiReg		3	.0191	.1155	33.908	23.17	< .0001
4	Coll_Math &Stat		4	.0118	.1274	21.198	14.49	.0001
5	Value		5	.0084	.1357	12.800	10.33	.0013
6	Amotivation		6	.0054	.1412	8.021	6.77	.0094
7	Required		7	.0030	.1442	6.241	3.79	.0519

The final multiple regression model I wanted to analyze was a model that incorporated the five demographic variables, alongside the motivation and statistics attitude subscales, and previously discussed interactions between them.

The resulting model was once again highly significant ($F(8, 1085) = 25.13$, $p < .0001$) with the R^2 for the final model rising to .1563 and a RMSE of 3.759. A table of parameter estimates (Table 19) and a summary of the stepwise selection process (Table 20) are again included. To refine the model, I added in the main effects for variables that were present in any significant interactions, even if the main effect itself was not significant. Any interactions that subsequently became non-significant (both the interactions between Identified Regulation and Amotivation with $p = .5002$, and between Intrinsic Motivation and Difficulty with $p = .4891$) were dropped from this final model. Additionally, after these interactions were dropped and the resulting model analyzed, both Difficulty ($p = .2683$) and Intrinsic Motivation ($p = .2072$) were also able to be removed, since they were initially only included based on their presence in since removed interactions. This new model (Table 21) has an R^2 of .1543 and is once again highly significant ($F(9, 1084) = 21.98$, $p < .0001$) with a RMSE of 3.765.

Table 19

Table of Parameter Estimates for Stepwise Regression Model Predicting CAOS Score from Motivation and Statistics Attitudes Subscales and Interactions, as well as Demographic Characteristics

Variable	b_i^*	*SE*	Type II SS	*F*	*p*
Intercept	18.59630	0.89760	6058.62125	429.23	< .0001
Female	-1.37021	0.25912	394.68572	27.96	< .0001
Required	-0.68670	0.34917	54.59388	3.87	.0495
Coll_Math&Stat	-0.39345	0.11188	174.551	12.37	.0005
CogComp	1.41431	0.16820	997.92049	70.70	< .0001
Intrinsic*Difficulty	0.04008	0.02235	45.38342	3.22	.0732
Amotivation*IdentiReg	-0.05839	0.02109	108.14114	7.66	.0057
IdentiReg*CogComp	-0.16542	0.03028	421.14821	29.84	< .0001
IdentiReg*Value	0.07764	0.02337	155.71998	11.03	.0009

Table 20

Summary of Stepwise Selection Process for Predicting CAOS Score from Motivation and Statistics Attitudes Subscales and Interactions, as well as Demographic Characteristics

Step	Variable Entered	Variable Removed	Number Vars In	Partial R^2	Model R^2	*C(p)*	*F*	*p*
1	CogComp		1	.0689	.0689	108.214	79.46	< .0001
2	IdentiReg* CogComp		2	.0299	.0989	72.310	35.60	< .0001
3	Female		3	.0198	.1187	49.237	24.06	< .0001
4	Coll_Math &Stats		4	.0117	.1304	36.424	14.39	.0002
5	IdentiReg* Value		5	.0107	.1411	24.834	13.35	.0003
6	Amotivation *IdentiReg		6	.0056	.1467	19.704	7.05	.0081
7	Required		7	.0033	.1500	17.567	4.10	.0431
8	Intrinsic* Difficulty		8	.0026	.1525	16.329	3.22	.0732

Table 21

Table of Parameter Estimates for Final Regression Model Predicting CAOS Score from Motivation and Statistics Attitudes Subscales and Interactions, as well as Demographic Characteristics

Variable	b_i^*	*SE*	*t*	*p*
Intercept	19.22754	2.64458	7.27	< .0001
Female	-1.36568	0.25661	-5.32	< .0001
Required	-0.79329	0.34549	-2.30	.0219
Coll_Math&Stats	-0.42368	0.10979	-3.86	.0001
IdentiReg	-0.07523	0.47949	-0.16	.8754
Amotivation	-0.27226	0.10550	-2.58	.0100
CogComp	1.84066	0.45075	4.08	< .0001
Value	-0.36264	0.44344	-0.82	.4137
IdentiReg*CogComp	-0.22539	0.08269	-2.73	.0065
IdentiReg*Value	0.14612	0.08315	1.76	.0791

Analysis 4. Item Analysis to Compare Conceptual Understanding of Students with Different Extremes of Motivational Traits

For the final part of my analysis, I performed an item analysis to look at which CAOS items were most affected by particular motivation traits. To do this I took the four different dimensions of motivation; Intrinsic Motivation, Amotivation, External Regulation, and Identified Regulation, and partitioned each scale into low, moderate, and high scores. For the purpose of this analysis, a low score was defined as a score in the lowest 27% on that dimension; a high score was one that placed in the largest 27% on that dimension. The 27% cutoff is considered optimal in item analysis (Kelley, 1939). In the event of a large number of equal values falling at these cutoffs, the cutoff was chosen to give as close to 27% as possible in these low and high scoring groups. All other observations were placed in the moderate group.

A two-sample comparison of proportions was performed for each CAOS item, comparing the proportion that answered the item correct out of those students with a low score on that dimension of motivation, with the proportion correct out of those students with a high score. This was the approach I chose because, unlike logistic regression, this technique doesn't require the relationship between the outcome and the trait to be linear. Additionally, categorizing the students as low, moderate, or high scoring on particular subscales allows a much more direct comparison of students with extremes on each trait. This comparison can be made both in terms of the percentage of each type of student that got each item correct, but also by comparing which incorrect answers were more successful distractor options for low and high scoring students on each trait.

As desirable as **intrinsic motivation** seems as a trait, it is perhaps surprising that students that scored highest on this dimension scored lower on the CAOS test than the low intrinsic motivation students, although the difference was not significant (p = .2293). Examination of the individual items revealed that more intrinsically motivated students performed worse on 26 of the 40 CAOS items, although most of the differences were not statistically significant. In comparing the items where the students more clearly differ, it appears that the more intrinsically motivated students did worse with items involving understanding of statistics theory and principles (distributions, law of large numbers, standard error) and better with some elements of the application of statistics (extrapolation, interpreting standard deviations, comparing boxplots and histograms). A detailed breakdown of the more pronounced differences is provided in Table 22.

In analyzing the effect of **amotivation** on conceptual understanding, of the 1148 students that completed both the AMS and the CAOS test, 46% received the lowest possible amotivation score of 4-out-of-28, meaning that the lower group in our analysis of amotivation was much larger here than the intended 27%. Those students who exhibited high amotivation (that is, students that were unsure of their abilities and doubting their decision to attend college) performed significantly worse on the CAOS test (p = .0430). Whereas this result is somewhat expected, it must also be stated that highly amotivated students actually scored higher than low amotivated students on 19-of-the-40 items on the test. Additionally, the students with moderate amotivation scores (i.e., those that were not in either of these two groups) actually outscored both groups, with a mean of 22.032 being significantly higher than the average score in the high amotivation group (p = .0182) but the difference was not significant when compared to the low amotivation group (p = .5318).

In a comparison of the items showing the greatest discrimination between groups, it appears that amotivated students have an overly-simplified view of statistics – being significantly more likely to think of the *p*-value as the probability of a treatment being effective in a drug trial (Item 11, $p < .0001$), and being more likely to struggle with comparisons of distributions. Many of the items that highly amotivated students did better on were the lower scoring items, suggesting that this overly-simplified approach could actually help these students on more difficult problems, or at least indicating that they are less inclined to see the same temptation in a strong distractor option. For example, a larger number of highly amotivated students may have been prepared to correctly make a causal inference in an experimental design with random assignment (Item 24) due to these students actually being less knowledgeable about when it is not appropriate to make a causal inference. The items with the largest group differences are indicated in Table 23.

In examining the effect of **external regulation** on conceptual understanding, students with the highest levels of external regulation (those students regulating their behavior to achieve positive, often job or salary related, outcomes) performed significantly worse than those students exhibiting low levels of external regulation ($p = .0039$), scoring lower on 28 of the 40 CAOS items. On examining items where the performance differences were more pronounced, it seems that externally regulated students are superior at spotting some of the traditional introductory statistics course "traps" (danger of extrapolation and outliers in regression, confidence interval misinterpretation), yet these same students struggle when it comes to applying statistics knowledge, particularly in more complex, multi-stage problems.

On item 6, where students were asked to graphically display a series of baseball batting averages to show the shape, center and spread of the distribution of the proportions, externally regulated students were much more likely to not only keep the player name on the horizontal axis, but also to reorder the players so the shape appeared normal. On item 34, externally regulated students were also much more likely to incorrectly believe that a sample of 500 values taken from a skewed population would have a normal distribution. These errors suggest that externally regulating students have a tendency to try and apply statistics rules and patterns to problems, even in situations where said application is neither appropriate nor correct.

A breakdown of the items where students with low and high external regulation differed most is included in Table 24.

The analysis of the conceptual statistical understanding of students with high levels of **identified regulation** (that is, students that attend college out of choice, yet do so to gain some personal value or satisfaction) reveals that these students score significantly worse than students with low levels of identified regulation ($p < .0001$). Whereas identified regulation, like intrinsic motivation earlier, seems like an obviously desirable trait, these results suggest that these traits actually have a negative effect on CAOS scores, with identified regulating students performing worse on 30 of the 40 items.

One possible explanation for this surprising result might be that an identified regulating student could be seeing personal value from college attendance precisely because of their own academic struggles, either prior to college or early in their college career. An examination of the CAOS items where identified regulating students outperformed those with low values of identified regulation adds weight to this argument, with identified regulating students being more likely to make elementary mistakes involving items typically introduced in the initial weeks of most introductory statistics courses, including boxplots (item 10), histograms (items 5 & 6) and standard deviations (item 14). Additionally, a failure to understand that statistics from small samples would vary more than statistics from large samples (item 16) suggests that students with large amounts of identified regulation are more likely to have substantial “knowledge gaps” that can ultimately hinder their performance throughout an introductory statistics course.

Table 25 contains a breakdown of the items where students with low and high amounts of identified regulation significantly differed on the percentage getting the item correct.

Table 22

Comparison of CAOS Item Performance between Low and High Intrinsically Motivated Students

Item	Measured Learning Outcome	Low Intrinsic Motivation	High Intrinsic Motivation	*p*
	Overall CAOS Score	21.899	21.495	.2293
CAOS items on which students with greater intrinsic motivation performed better:				
		Proportion of Students Correct		
25	Ability to recognize a correct interpretation of a *p*-value.	.413	.544	.0012
39	Understanding of when it is not wise to extrapolate using a regression model.	.060	.111	.0272
15	Ability to correctly estimate standard deviations for different histograms. Understands highest standard deviation would be for a graph with the most spread (typically) away from the center.	.584	.635	.1960
2	Ability to recognize two different graphical representations of the same data (boxplot and histogram).	.470	.521	.2064
CAOS items on which students with lower intrinsic motivation performed better:				
		Proportion of Students Correct		
11	Ability to compare groups by considering where most of the data are, and focusing on distributions as single entities.	.973	.922	.0048
34	Understanding of the law of large numbers for a large sample by selecting an appropriate sample from a population given the sample size.	.644	.554	.0231
16	Understanding that statistics from small samples vary more than statistics from large samples.	.413	.339	.0602
31	Ability to correctly interpret a confidence interval.	.849	.801	.1229

Based on sample sizes of 298 in low intrinsic motivation group and 307 in high intrinsic motivation group.

Table 23

Comparison of CAOS Item Performance between Low and High Amotivation Students

Item	Measured Learning Outcome	Low Amotivation	High Amotivation	*p*
	Overall CAOS Score	21.850	21.256	.0430
CAOS items on which students with greater amotivation performed better:				
		Proportion of Students Correct		
32	Understanding of how sampling error is used to make an informal inference about a sample mean.	.065	.105	.0343
7	Understanding of the purpose of randomization in an experiment.	.032	.054	.1168
10	Understanding of the interpretation of a median in the context of boxplots.	.203	.240	.2134
24	Understanding that an experimental design with random assignment supports causal inference.	.641	.681	.2482
CAOS items on which students with lower amotivation performed better:				
		Proportion of Students Correct		
27	Ability to recognize an incorrect interpretation of a *p*-value (probability that a treatment is effective).	.562	.419	<.0001
12	Ability to compare groups by comparing differences in averages.	.922	.866	.0081
11	Ability to compare groups by considering where most of the data are, and focusing on distributions as single entities.	.953	.914	.0240
19	Understanding that low *p*-values are desirable in research studies.	.740	.671	.0320

Based on sample sizes of 527 in low amotivation group and 313 in high amotivation group.

Table 24

Comparison of CAOS Item Performance between Low and High External Regulation Students

Item	Measured Learning Outcome	Low External Regulation	High External Regulation	*p*
	Overall CAOS Score	22.035	21.082	.0039
CAOS items on which students with greater external regulation performed better:				
		Proportion of Students Correct		
15	Ability to correctly estimate standard deviations for different histograms. Understands highest standard deviation would be for a graph with the most spread (typically) away from the center.	.540	.621	.0456
21	Ability to correctly describe a bivariate relationship shown in a scatterplot when there is an outlier (influential point).	.786	.833	.1332
39	Understanding of when it is not wise to extrapolate using a regression model.	.060	.085	.2381
30	Ability to detect a misinterpretation of a confidence level (percentage of all possible sample means between confidence limits).	.312	.351	.3122
CAOS items on which students with lower external regulation performed better:				
		Proportion of Students Correct		
6	Understanding to properly describe the distribution of a quantitative variable, need a graph like a histogram that places the variable along the horizontal axis and frequency along the vertical axis.	.214	.103	.0002
34	Understanding of the law of large numbers for a large sample by selecting an appropriate sample from a population given the sample size.	.649	.542	.0077
37	Understanding of how to simulate data to find the probability of an observed value.	.158	.091	.0122
10	Understanding of the interpretation of a median in the context of boxplots.	.242	.169	.0265

Based on sample sizes of 285 in low external regulation group and 319 in high external regulation group.

Table 25

Comparison of CAOS Item Performance between Low and High Identified Regulation Students

Item	Measured Learning Outcome	Low Identified Regulation	High Identified Regulation	*p*
	Overall CAOS Score	22.319	21.016	<.0001
CAOS items on which students with greater identified regulation performed better:				
		Proportion of Students Correct		
39	Understanding of when it is not wise to extrapolate using a regression model.	.070	.117	.0369
30	Ability to detect a misinterpretation of a confidence level (percentage of all possible sample means between confidence limits).	.319	.394	.0449
2	Ability to recognize two different graphical representations of the same data (boxplot and histogram).	.497	.562	.0963
CAOS items on which students with lower identified regulation performed better:				
		Proportion of Students Correct		
16	Understanding that statistics from small samples vary more than statistics from large samples.	.436	.292	.0001
10	Understanding of the interpretation of a median in the context of boxplots.	.263	.159	.0011
6	Understanding to properly describe the distribution of a quantitative variable, need a graph like a histogram that places the variable along the horizontal axis and frequency along the vertical axis.	.216	.127	.0025
36	Understanding of how to calculate appropriate ratios to find conditional probabilities using a table of data.	.579	.483	.0133
35	Understanding of how to select an appropriate sampling distribution for a particular population and sample size.	.512	.416	.0139
34	Understanding of the law of large numbers for a large sample by selecting an appropriate sample from a population given the sample size.	.643	.549	.0140
37	Understanding of how to simulate data to find the probability of an observed value.	.167	.102	.0149

5	Ability to visualize and match a histogram to a description of a variable (uniform distribution for the last digit of phone numbers sampled from a phone book).	.810	.733	.0191
14	Ability to correctly estimate and compare standard deviations for different histograms. Understands lowest standard deviation would be for a graph with the least spread (typically) away from the center.	.611	.527	.0295
28	Ability to detect a misinterpretation of a confidence level (the percentage of sample data between confidence limits).	.456	.375	.0342

Based on sample sizes of 285 in low identified regulation group and 319 in high identified regulation group.

CHAPTER 5

SUMMARY AND DISCUSSION

Summary of the Study

This chapter summarizes the main findings of this study, as well as further implications in statistics teaching.

The first part of this study attempted to find the optimal configuration of the Academic Motivation Scale (AMS) through principal component analysis. A four subscale configuration with some items removed was deemed most appropriate, mirroring the configuration identified by Smith et al. (2012). These four subscales – Intrinsic Motivation, Amotivation, External Regulation, and Identified Regulation – explained 62.3% of the variability found in the data, through the use of just 17-of-the-28 items.

The main distinctions between this four subscale solution and the original seven-subscale configuration identified by Vallerand et al. (1992) is the collapsing of the three distinct, unordered subscales of intrinsic motivation (intrinsic motivation to know, to accomplish things, and to experience stimulation) into one single subscale, as well as omitting introjected regulation from the three ordered subscales of extrinsic motivation (external, introjected and identified regulation). Notably, all four intrinsic motivation to accomplish items were removed due to cross loading on both the Intrinsic Motivation and Identified Regulation subscales, lending credence to the argument made by Fairchild et al. (2005) that intrinsic and extrinsic motivation may not be as mutually exclusive as suggested by Self-Determination Theory (SDT).

SDT also states the continuum of extrinsic motivation subscales should be arranged with external regulation being followed by introjected regulation, which in turn would be followed by identified regulation. Smith et al. (2012) had previously found that items measuring the two extremes of this supposed continuum (external and identified regulation) loaded on the one subscale, with the introjected regulation items loading on another. This finding was supported by Fairchild et al. (2005), who noted that the original external and identified regulation items all focus on career and job aspirations, whereas the introjected regulation items focus instead on aspects of competency and self-worth, although the arrangement is contrary to SDT. After the omission of the original identified regulation items from this model due to significant cross-loadings, the model derived here has more distinct subscales that are more consistent with SDT.

After analysis of the correlations between motivation subscales, the arrangement of subscales is also consistent with SDT, with a continuum of subscales following the taxonomy of human motivations proposed by Ryan & Deci (2000). This continuum, complete with corresponding correlations between subscales, is shown in Figure 2 and is arranged based on the extent to which the motivation for the behavior – in this case college attendance – comes from the student. Note that the correlations depicted are those between average score on the subscale items, rather than between factor scores from the SEM model, hence the differences between the correlations depicted here and in Figure 1.

Analysis of the correlations between statistics attitude averages reveal a pattern similar to that identified by Schau et al. (1995). These correlations are shown in Figure 3. The correlations between CAOS scores and all eight subscales averages are shown in Figure 4.

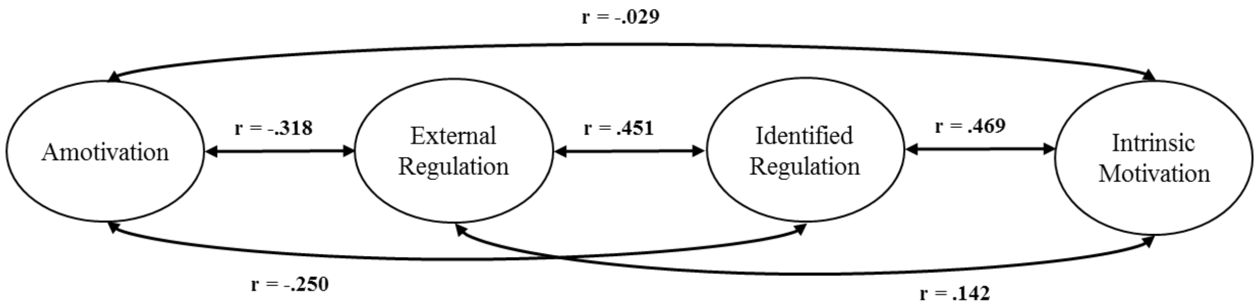

Figure 2

Arrangement of and correlations between motivation subscales consistent with Ryan & Deci (2000) taxonomy of human motivations.

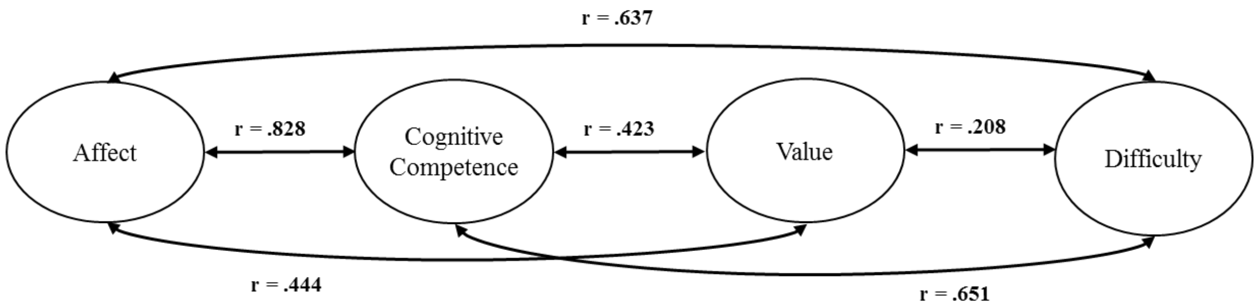

Figure 3

Correlation between statistics attitude subscales.

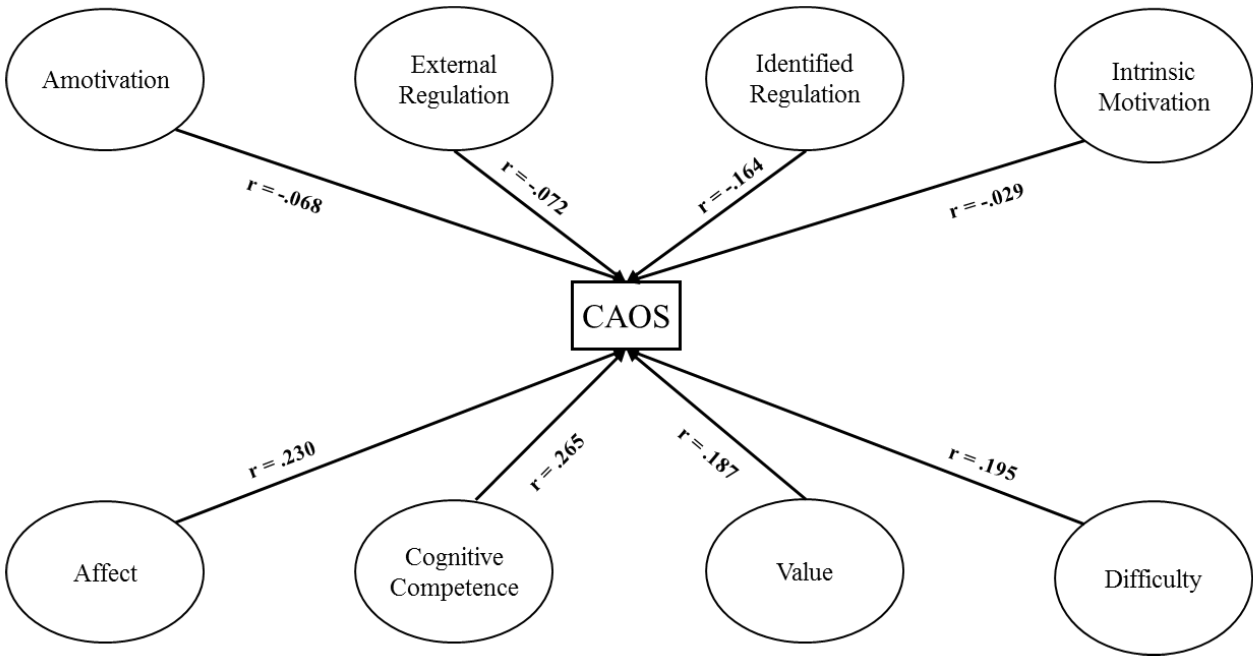

Figure 4

Correlation between subscales and CAOS scores.

Analysis of the correlations between the motivation subscales and CAOS scores (Appendix F) reveals the strongest correlation between CAOS scores and Identified Regulation ($r = -.164$, $p < .001$) followed by External Regulation ($r = -.072$, $p = .018$) and Amotivation ($r = -.068$, $p = .025$). Intrinsic Motivation ($r = -.029$, $p = .336$) was the only motivation subscale not significantly correlated with CAOS score. Similar results were found through multiple regression analysis, where both Identified Regulation and Amotivation (though not External Regulation) were present in the final multiple regression model predicting CAOS scores, although the former was only present in the form of interactions with Cognitive Competence and Value (Table 21).

The analysis comparing students with high and low scores on motivational traits also found that students with high values on the identified regulation subscale performed significantly worse than students with low values on this subscale ($p < .0001$, Table 25), with highly amotivated students also performing worse than students with low amounts of amotivation ($p = .0430$, Table 23). There were also significant differences between students with high and low levels of external regulation, with students with large amounts of external regulation averaging lower on the CAOS test ($p = .0039$, Table 24) by almost a point.

The above suggests that the most important motivational traits in determining a student's conceptual understanding of statistics are identified regulation, amotivation and external regulation, with low scoring students on these traits having greater conceptual understanding. Since the identified regulation subscale contains items such as, "because I want to show myself that I can succeed in my studies" and, "to prove to myself that I am capable of completing my college degree," this would suggest that students who are not already confident enough in their ability to succeed in college struggle most in gaining a conceptual understanding in the course.

The interaction between identified regulation and both cognitive competence and value illustrate how a lack of an inherent belief in self and a lack of worth or usefulness seen in statistics in general can combine to the detriment of a student's ultimate conceptual understanding of statistics.

The significance of the amotivation trait suggests that highly amotivated students – in other words, students that feel like they are wasting their time in school and question why they are continuing their education – do not gain as strong a conceptual understanding of statistics either. However, the results indicate that the relationship between amotivation and conceptual understanding is not linear, with some small amount of amotivation potentially being a positive.

The negative correlation between external regulation and CAOS score indicates that students attending college based on externally regulated (typically career-related) goals gain significantly less conceptual understanding of statistics. These students tend to spend more time focusing on spotting traps and learning rules-of-thumb, and less time developing a true understanding of the subject matter.

A similar analysis of the statistics attitudes subscales finds that all four attitude subscales are significantly correlated with CAOS score (Appendix F), however only Cognitive Competence and Value were present in the final multiple regression model shown in Table 21. These results suggest that students who are confident in their ability to do well at statistics and that see value in the topic will gain a stronger conceptual understanding of statistics from the course.

To examine the effect of the various demographic characteristics on CAOS scores, Table 21 shows that gender, whether the course was required, and the number of college mathematics and statistics courses previously taken were all significant predictors, whereas student class

(Freshman/ Sophomore/ Junior/ Senior) and whether the student was age 23-or-older were not significant. Further analysis of the parameter estimates in this table allows the following conclusions to be made about scores on the 40 question CAOS 4 test. Note that all these conclusions can only be made **at fixed levels of the other demographic, attitudinal and motivational subscales**:

- Females can be expected to score 1.36 points lower than males
- Those required to take the course can be expected to score 0.79 lower than those opting to take it
- Every additional college course in mathematics and statistics the student has taken (up-to-four) would lead to a prediction of CAOS score that is .424 lower
- For every additional point higher the students' amotivation average is (out of 7, with higher scores indicating a more amotivated student that is more frustrated with college), their CAOS score (out of 40) would be predicted to be 0.272 lower.
- For every additional point higher the students' identified regulation average is (out of 7, with higher scores more indicative of a student identifying college attendance as important and doing it out of choice), we would expect their CAOS score (out of 40) to decrease by 0.075, with a further decrease of 0.225 times their cognitive comprehension average (out of 7, higher scores indicating the student possessing a more positive attitudes about their intellectual knowledge and skills when applied to statistics), with an offsetting increase of 0.146 times their value average (out of 7, higher scores indicating a more positive attitude towards the usefulness, relevance, and worth of statistics to their life).

- For every additional point higher the students' cognitive comprehension average is, we would expect their CAOS score to increase by 1.841, with an offsetting decrease of 0.225 times their identified regulation average.
- For every additional point higher the students' value average is, we would expect their CAOS score to decrease by 0.363, with an offsetting increase of 0.146 times their identified regulation score.

Limitations and Implications for Future Research

One limitation of this study is that conceptual understanding was measured only at the end of the course, rather than following a pre-test/post-test format. This means that, rather than measuring the conceptual understanding that the student *gains* from the course, it instead measures the conceptual statistics understanding that the student *has* at the end of the course. The study was set up this way partly for practical reasons (the students already needed to complete two survey instruments at the start of the course), but also because maximizing conceptual understanding at the end of the course (rather than maximizing gains in conceptual understanding) seems more congruent with the aims of statistics education. However, when noting the positive correlation between cognitive competence and CAOS score, for example, this does not necessarily mean that students with greater cognitive competence *gain* a greater conceptual understanding of statistics from the course, since some-or-all of their greater understanding might be due to their starting the course with a greater competence towards conceptual ideas in the first place.

A further limitation of the study is that students' motivations for attending college do not necessarily coincide with their motivations for taking a statistics course. A student that is intrinsically motivated to attend college could still be taking statistics purely as a requirement, and so see little inherent purpose in statistics. The use of a variable measuring whether the course is required for a student's major was intended to offset this effect, yet a different motivational scale, measuring a student's motivation for taking a course, may be more appropriate than the AMS. Also, there are a number of items on the AMS measuring a student's intrinsic motivation by asking about pleasure gained from reading different authors, yet no items concerning pleasure derived from "solving a difficult puzzle" or "reasoning out a solution to a problem". The development of an instrument measuring a student's motivation for taking a course, that is equally as suitable for qualitative and quantitative courses, would certainly be an area suitable future research.

Teaching Implications

In the analysis comparing the CAOS scores of students with low external regulation against those with high values on the external regulation subscale, students with high values scored lower (p = .0039) by almost a whole point on average (Table 24). This suggests that students taking the course to improve job prospects or potential salary do not achieve as strong a conceptual statistical understanding as students without this objective. Applying statistical techniques to real world problems, and emphasizing where these techniques and methods are used in a variety of fields and occupations, would hopefully motivate these externally regulated students to connect to the material more completely. This use of real world application would also likely help the highly amotivated student see purpose in statistics, and therefore motivate a more meaningful understanding.

Similarly, students with large values on the identified regulation subscale scored lower than those students with small values ($p < .0001$) by an average of 1.3 points. This was the most significant motivational subscale when comparing groups of students, and suggests that students attending college to show themselves that they are intelligent and that they're capable of completing a college degree do struggle to achieve the same level of conceptual understanding of someone without that objective. Emphasizing the logic behind the process, and drawing parallels between techniques and pre-existing thought processes the student may already have, would hopefully work to deconstruct any ideas the low identified regulation student might have about the "difficulty" of statistics and instead help them see the logical underpinnings of the techniques.

Conclusion

The evidence confirms a four factor structure to the Academic Motivation Scale is appropriate. Of those four factors, identified regulation, amotivation and external regulation are most strongly linked (and negatively so) with a student's conceptual understanding of statistics. These factors, along with the statistics attitude factors of cognitive competence and the value that a student sees in statistics are all significant factors in determining a student's level of conceptual understanding.

Based on these conclusions, it is suggested that this study contributes to statistics education research in the following ways: (1) by providing further justification for the four-factor motivational structure suggested by Smith et al. (2012); (2) by identifying which motivational and attitudinal factors most affect a student's conceptual understanding of statistics; and (3) by identifying which type of mistakes a student with particular motivational traits might be more likely to make, and which misconception they are more likely to have. Identifying these traits, and their associated misconceptions, are important contributions to the field of statistics education.

REFERENCES

Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. (1978). Learned helplessness in humans: critique and reformulation. *Journal of Abnormal Psychology, 87*, 49-74.

Albarracin, D., Johnson, B. T., & Zanna, M. (Eds.). (2005). *The handbook of attitudes*. Mahwah, NJ: Lawrence Erlbaum

Artino, A. R. **(**2005). *Review of the motivated strategies for learning questionnaire*. (ERIC Document Reproduction Service No. ED499083)

Benson, J. (1989). Structural components of statistical test anxiety in adults: An exploratory model. *The Journal of Experimental Education, 57*, 247-261.

Benware, C., & Deci, E. L. (1984). The quality of learning with an active versus passive motivational set. *American Educational Research Journal*, *21*, 755-766.

Boslaugh, S., & Watters, P. A. (2008). *Statistics in a nutshell*. Sebastopol, CA: O'Reilly Media, Inc.

Cashin, S. E., & Elmore, P. B. (2005). The survey of attitudes toward statistics scale: A construct validity study. *Educational and Psychological Measurement, 65*, 509-524.

Chance, B. L. (2002). Components of statistical thinking and implications for instruction and assessment. *Journal of Statistics Education [Online]*, *10*(3). www.amstat.org/publications/jse/v10n3/chance.html

Cobb, G. (1992). Teaching statistics. In *Heeding the Call for Change: Suggestions for Currricular Action, MAA Notes, 22*, 3-33.

Cokley, K. O. (2000). Examining the validity of the academic motivation scale by comparing scale construction to self-determination theory. *Psychological Reports, 86*, 560–564.

Cokley, K. O., Bernard, N., Cunningham, D., & Motoike, J. (2001). A psychometric investigation of the academic motivation scale using a United States sample. *Measurement and Evaluation in Counseling and Development, 34*, 109–119.

Cruise J. R., Cash R. W., & Bolton L. D. (1985). Development and validation of an instrument to measure statistical anxiety. *American Statistical Association Proceedings of the Section on Statistical Education* (pp. 92-97). Alexandria, VA: American Statistical Association.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1*, 16-29.

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

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*, 227–268.

delMas, R. C. (2002). Statistical literacy, reasoning, and learning: A commentary. *Journal of Statistics Education [Online], 10*(3). http://www.amstat.org/publications/jse/v10n3/delmas_discussion.html

delMas, R., Garfield, J., Ooms, A., & Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal, 6*(2), 28-58.

Dempster, M., & McCorry, N. K. (2009). The role of previous experience and attitudes toward statistics in statistics assessment outcomes among undergraduate psychology students. *Journal of Statistics Education*, *17*(2).

Emmioğlu , E., & Capa-Aydin, Y. (2012). Attitudes and achievement in statistics: A meta-analysis study. *Statistics Education Research Journal*, *11*(2), 95-102. Retrieved from www.stat.auckland.ac.nz/serj.

Fairchild A. J., Horst S. J., Finney S. J., & Barron K. E. (2005). Evaluating existing and new validity evidence for the academic motivation scale. *Contemporary Educational Psychology, 30,* 331-358.

Fitzgerald, S. M., Jurs, S., & Hudson, L. M. (1996) A model predicting statistics achievement among graduate students. *College Student Journal*, *30*, 361–366.

Fortier, M. S., Vallerand, R. J., & Guay, F. (1995). Academic motivation and school performance: Toward a structural model. *Contemporary Educational Psychology, 20,* 257-274.

Friel, S., Bright, G., Frierson, D., & Kader, G. (1997). A framework for assessing knowledge and learning in statistics (K-8) in I. Gal & J. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 55-63). Burke, VA: IOS Press.

GAISE. (2006). *College Report of the Guidelines for Assessment and Instruction in Statistics Education Project* downloaded from http://www.amstat.org/education/gaise/GAISECollege.htm

Gal, I. (Ed.). (2000). *Adult numeracy development: Theory, research, practice*, Cresskill, NJ: Hampton Press.

Gal, I., Ginsburg, L., & Schau, C. (1997). Monitoring attitudes and beliefs in statistics education. In I. Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 37–51). Amsterdam, Netherlands: IOS Press. [Online: http://www.stat.auckland.ac.nz/~iase/publications/assessbk/chapter04.pdf.]

Garfield, J. (1998). The statistical reasoning assessment: Development and validation of a research tool. *Proceedings of the Fifth International Conference on Teaching Statistics*. In L. Pereira-Mendoza (Ed.), Voorburg, The Netherlands: International Statistical Institute (pp. 781-786).

Garfield, J. (1999). Thinking about statistical reasoning, thinking, and literacy. Paper presented at *First Annual Roundtable on Statistical Thinking, Reasoning, and Literacy (STRL-1)*.

Garfield, J. (2002). The challenge of developing statistical reasoning. *Journal of Statistics Education [Online], 10*(3). www.amstat.org/publications/jse/v10n3/garfield.html

Garfield, J. (2003). Assessing statistical reasoning. *Statistics Education Research Journal, 2*(1), 22-38. [Online: www.stat.auckland.ac.nz/~iase/serj/SERJ2 (1).pdf]

Garfield, J., delMas, R., & Chance, B. (2002). *The Assessment Resource Tools for Improving Statistical Thinking (ARTIST) Project.* NSF CCLI grant ASA- 0206571. [Online: https://apps3.cehd.umn.edu/artist/caos.html]

Gottfried, A. E. (1990). Academic intrinsic motivation in young elementary school children. *Journal of Educational Psychology, 82*, 525-538.

Hawkins, A., Jolliffe, F., & Glickman, L. (1992). *Teaching statistical concepts*, London, England: Longman Publishers.

Hirsch, L., & O'Donnell, A. (2001). Representativeness in statistical reasoning: Identifying and assessing misconceptions. *Journal of Statistics Education [Online], 9*(2)

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods, 6*(1), 53-60.

Hu, L., Bentler, P. M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin, 112*, 351-362.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55.

Judi, H. M., Ashaari, N. S., Mohamed, H., & Wook, T. M. T. (2011). Students profile based on attitude towards statistics. *Procedia Social and Behavioral Sciences, 18*, 266–272

Kalinowski, P. (2010). Identifying misconceptions about confidence intervals. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics* [CD-ROM]. Ljubljana, Slovenia: IASE, July 2010. www.stat.auckland.ac.nz/~iase/publications.php

Kelley, T. L. (1939). The selection of upper and lower groups for the validation of test items. *Journal of Educational Psychology, 30*, 17-24.

Komarraju M., Karau S. J., & Schmeck, R. R. (2009). Role of the big five personality traits in predicting college students' academic motivation and achievement. *Learning and Individual Differences, 19*, 47–52.

Konold, C. (1995). Issues in assessing conceptual understanding in probability and statistics. *Journal of Statistics Education [Online], 3*(1). http://www.amstat.org/publications/jse/v3n1/konold.html

Linnenbrink, E., & Pintrich, P. (2002). The role of motivational beliefs in conceptual change. In M. Limon & L. Mason (Eds.) *Reconsidering conceptual change: Issues in theory and practice* (pp. 115-135). Dordecht, The Netherlands: Kluwer Academic. doi:10.1007/0-306-47637-1_6

Lovett, M. (2001). A collaborative convergence on studying reasoning processes: A case study in statistics. In D. Klahr & S. Carver (Eds.) *Cognition and instruction: Twenty-five years of progress* (pp. 347-384). Mahwah, NJ: Lawrence Erlbaum

Mathews, D., & Clark, J. (2003). *Successful students' conceptions of mean, standard deviation and the central limit theorem*. Unpublished paper. Retrieved March 26, 2012, from http://www1.hollins.edu/faculty/clarkjm/stats1.pdf

Mji, A., & Onwuegbuzie, A. (2004). Evidence of score reliability and validity of the statistical anxiety rating scale among technikon students in South Africa. *Measurement and Evaluation in Counselling and Development, 36*, 238-251.

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

National Research Council. (1991). *Moving beyond myths: Revitalizing undergraduate mathematics*. Washington, DC: National Academy Press.

Onwuegbuzie, A. J., Slate, J., Paterson, F., Watson, M., & Schwartz, R. (2000). Factors associated with underachievement in educational research courses. *Research in the Schools, 7*, 53–65.

Pantziara, M., & Philippou, G. (2006). Measuring and relating primary students' motives, goals and performance in mathematics (pp. 321-328). In J. Novotná, H. Moraová, M. Krátká, & N. Stehlíková (Eds.), *Proceedings 30th Conference of the International Group for the Psychology of Mathematics Education, 4* (pp. 321-328). Prague, Czech Republic: PME.

Patrick, H., Anderman, L. H., Ryan, A. M., Edelin, K., & Midgley, C. (2001). Teachers' communication of goal orientations in four fifth-grade classrooms. *The Elementary School Journal, 102*, 35-58.

Pelletier, L. G., Fortier, M. S., Vallerand, R. J., Tuson, K. M., Briere, N. M., & Blais, M. R. (1995). Toward a new measure of intrinsic motivation, extrinsic motivation, and amotivation in sports: The sport motivation scale. *Journal of Sport & Exercise Psychology*, *17*, 35-53.

Pintrich, R. R., & DeGroot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance, *Journal of Educational Psychology, 82*, 33-40.

Reed-Rhoads, T., Murphy, T. J., & Terry, R. (2006). The Statistics Concept Inventory (SCI). [*Online:* http://coecs.ou.edu/sci/]

Roberts, D. M., & Bilderback, E. W. (1980). Reliability and validity of a statistics attitude survey. *Educational and Psychological Measurement*, *40*, 235-238.

Roberts, D. M., & Reese, C. M. (1987). A comparison of two scales measuring attitudes towards statistics. *Educational and Psychological Measurement*, *47*, 759-764.

Roberts, D. M., & Saxe, J. E. (1982). Validity of a statistics attitude survey – a follow-up study. *Educational and Psychological Measurement*, *42*, 907-912.

Rodriguez, R. (2012). More than 1 million and counting: The growth of Advanced Placement statistics. *AmStat News*. Retrieved October 8, 2012.

Ruggeri, K., Dempster, M., Hanna, D., & Cleary, C. (2008). Experiences and expectations: The real reason nobody likes stats. *Psychology Teaching Review*, *14(2)*, 75-83.

Rumsey, D. J. (2002). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education [Online], 10*(3). www.amstat.org/publications/jse/v10n3/rumsey2.html

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54-67.

Schau, C. (2003, August). Students' attitudes: The "other" important outcome in statistics education. Paper presented at the *Joint Statistical Meetings*, San Francisco, CA. [Online: http://evaluationandstatistics.com/JSM2003.pdf]

Schau, C., Stevens, J., Dauphinee, T. L., & Del Vecchio, A. (1995). The development and validation of the Survey of Attitudes Toward Statistics. *Educational and Psychological Measurement*, *55,* 868-875.

Schield, M. (1999). Statistical literacy: Thinking critically about statistics. *APDU of Significance, 1*(1). 15-20. [*Online:* www.StatLit.org/pdf/1999SchieldAPDU.pdf.]

Shia, R. (1998). *Academic intrinsic and extrinsic motivation and metacognition. Assessing academic intrinsic motivation: A look at student goals and personal strategy*. Wheeling Jesuit University. Retrieved June 8, 2005 from the web at: http://www.cet.edu/research/pdf/motivation.pdf .

Smith, K. J., Davy, J. A., & Rosenberg, D. L. (2010). An examination of the validity of the academic motivation scale with a United States business student sample. *Psychological Reports, 106*, 323-341.

Smith, K. J., Davy, J. A., & Rosenberg, D. L. (2012). An empirical analysis of an alternative configuration of the Academic Motivation Scale. *Assessment in Education: Principles, Policy & Practice*, *19*(2), 1-20.

Sotos, A., 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,* 98–113

Sotos, A., Vanhoof, S., Van den Noortgate, W., & Onghena, P. (2009). How confident are students in their misconceptions about hypothesis tests? *Journal of Statistics Education [Online]*, *17*(2). www.amstat.org/publications/jse/v17n2/castrosotos.html

Stone, A., Allen, K., Reed-Rhoads, T., Murphy, T. J., Shehab, R. L., & Saha C. (2003). The Statistics Concept Inventory: A pilot study. *ASEE/IEEE 33rd Annual Frontiers in Education Conference*, Boulder, CO. doi: 10.1109/FIE.2003.1263336

Utts, J. (2003). What educated citizens should know about statistics and probability. *The American Statistician*, *57*(2), 74-79.

Vallerand, R. J., & O'Connor, B. P. (1989). Motivation in the elderly: A theoretical framework and some promising findings. *Canadian Psychology, 30*, 538–550.

Vallerand, R. J., Blais, M. R., BriBre, N. M., & Pelletier, L. G. (1989). Construction et validation de 1'Echelle de Motivation en Education [Construction and validation of the Academic Motivation Scale]. *Canadian Journal of Behavioral Sciences*, *21*, 323-349.

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C. &.Vallieres, E. F. (1992). The Academic Motivation Scale: A measure of intrinsic, extrinsic and amotivation in education. *Educational and Psychological Measurement*, *52*, 1003–1017.

Wheaton, B., Muthen, B., Alwin, D., F., & Summers, G. (1977). Assessing reliability and stability in panel models. *Sociological Methodology, 8*, 84-136.

Wise, S. L. (1985), The development and validation of a scale measuring attitudes toward statistics. *Educational and Psychological Measurement*, *45*, 401-405.

APPENDICES

Appendix A: 40 Measured Learning Outcomes of CAOS Test Items

1. Ability to describe and interpret the overall distribution of a variable as displayed in a histogram, including referring to the context of the data.
2. Ability to recognize two different graphical representations of the same data (boxplot and histogram).
3. Ability to visualize and match a histogram to a description of a variable (neg. skewed distribution for scores on an easy quiz).
4. Ability to visualize and match a histogram to a description of a variable (bell-shaped distribution for wrist circumferences of newborn female infants).
5. Ability to visualize and match a histogram to a description of a variable (uniform distribution for the last digit of phone numbers sampled from a phone book).
6. Understanding to properly describe the distribution of a quantitative variable, need a graph like a histogram that places the variable along the horizontal axis and frequency along the vertical axis.
7. Understanding of the purpose of randomization in an experiment.
8. Ability to determine which of two boxplots represents a larger standard deviation.
9. Understanding that boxplots do not provide estimates for percentages of data above or below values except for the quartiles.
10. Understanding of the interpretation of a median in the context of boxplots.

11. Ability to compare groups by considering where most of the data are, and focusing on distributions as single entities.
12. Ability to compare groups by comparing differences in averages.
13. Understanding that comparing two groups does not require equal sample sizes in each group, especially if both sets of data are large.
14. Ability to correctly estimate and compare standard deviations for different histograms. Understands lowest standard deviation would be for a graph with the least spread (typically) away from the center.
15. Ability to correctly estimate standard deviations for different histograms. Understands highest standard deviation would be for a graph with the most spread (typically) away from the center.
16. Understanding that statistics from small samples vary more than statistics from large samples.
17. Understanding of expected patterns in sampling variability.
18. Understanding of the meaning of variability in the context of repeated measurements and in a context where small variability is desired.
19. Understanding that low p-values are desirable in research studies.
20. Ability to match a scatterplot to a verbal description of a bivariate relationship.
21. Ability to correctly describe a bivariate relationship shown in a scatterplot when there is an outlier (influential point).
22. Understanding that correlation does not imply causation.
23. Understanding that no statistical significance does not guarantee that there is no effect.

24. Understanding that an experimental design with random assignment supports causal inference.

25. Ability to recognize a correct interpretation of a p-value.

26. Ability to recognize an incorrect interpretation of a p-value (probability that a treatment is not effective).

27. Ability to recognize an incorrect interpretation of a p-value (probability that a treatment is effective).

28. Ability to detect a misinterpretation of a confidence level (the percentage of sample data between confidence limits)

29. Ability to detect a misinterpretation of a confidence level (percentage of population data values between confidence limits).

30. Ability to detect a misinterpretation of a confidence level (percentage of all possible sample means between confidence limits).

31. Ability to correctly interpret a confidence interval.

32. Understanding of how sampling error is used to make an informal inference about a sample mean.

33. Understanding that a distribution with the median larger than mean is most likely skewed to the left.

34. Understanding of the law of large numbers for a large sample by selecting an appropriate sample from a population given the sample size.

35. Understanding of how to select an appropriate sampling distribution for a particular population and sample size.

36. Understanding of how to calculate appropriate ratios to find conditional probabilities using a table of data.
37. Understanding of how to simulate data to find the probability of an observed value.
38. Understanding of the factors that allow a sample of data to be generalized to the population.
39. Understanding of when it is not wise to extrapolate using a regression model.
40. Understanding of the logic of a significance test when the null hypothesis is rejected.

Appendix B: 28 Items of the SATS-28 Survey

Respondents rated their level of agreement with each item on a scale from 1 (strongly disagree) to 7 (strongly agree). The dimension measured by the item is indicated in parentheses. An asterisk next to the item indicates a negatively worded item that was reverse coded in the analysis.

1. I will like statistics. (AFFECT 1)
2. I will feel insecure when I have to do statistics problems. (AFFECT 2*)
3. I will have trouble understanding statistics because of how I think. (COGNITIVE COMPETENCE 1*)
4. Statistics formulas are easy to understand. (DIFFICULTY 1)
5. Statistics is worthless. (VALUE 1*)
6. Statistics is a complicated subject. (DIFFICULTY 2*)
7. Statistics should be a required part of my professional training. (VALUE 2)
8. Statistical skills will make me more employable. (VALUE 3)
9. I will have no idea of what's going on in statistics. (COGNITIVE COMPETENCE 2*)
10. Statistics is not useful to the typical professional. (VALUE 4*)
11. I will get frustrated going over statistics tests in class. (AFFECT 3*)
12. Statistical thinking is not applicable in my life outside my job. (VALUE 5*)
13. I use statistics in my everyday life. (VALUE 6)
14. I will be under stress during statistics class. (AFFECT 4*)
15. I will enjoy taking statistics courses. (AFFECT 5)

16. Statistics conclusions are rarely presented in everyday life. (VALUE 7*)

17. Statistics is a subject quickly learned by most people. (DIFFICULTY 3)

18. Learning statistics requires a great deal of discipline. (DIFFICULTY 4*)

19. I will have no application for statistics in my profession. (VALUE 8*)

20. I will make a lot of math errors in statistics. (COGNITIVE COMPETENCE 3*)

21. I am scared by statistics. (AFFECT 6*)

22. Statistics involves massive computations. (DIFFICULTY 5*)

23. I can learn statistics. (COGNITIVE COMPETENCE 4)

24. I will understand statistics equations. (COGNITIVE COMPETENCE 5)

25. Statistics is irrelevant in my life. (VALUE 9*)

26. Statistics is highly technical. (DIFFICULTY 6*)

27. I will find it difficult to understand statistics concepts. (COGNITIVE COMPETENCE 6*)

28. Most people have to learn a new way of thinking to do statistics. (DIFFICULTY 7*)

Appendix C: Demographic Survey

29. What is your gender? (Male/ Female)

30. What is your year of college? (Freshman/ Sophomore/ Junior/ Senior/ Other)

31. Is this course required for your college degree? (Yes/ No)

32. Which of these age groups do you belong to? (17 or under/ 18-22/ 23-28/ 29-34/ 35+)

33. How many college mathematics and/or statistics courses have you completed, not counting this semester? (0/ 1/ 2/ 3/ 4+)

Appendix D: 28 Items of the AMS

Respondents will rate each item based on its level of correspondence to one of their reasons for going to college, on a scale from 1 (does not correspond at all) to 7 (corresponds exactly).

1. Because with only a high-school degree I would not find a high-paying job later on. (ER 1)
2. Because I experience pleasure and satisfaction while learning new things. (IMTK 1)
3. Because I think that a college education will help me better prepare for the career I have chosen. (IDENT 1)
4. For the intense feelings I experience when I am communicating my own ideas to others. (IMTS 1)
5. Honestly, I don't know; I really feel that I am wasting my time in school. (AMOT 1)
6. For the pleasure I experience while surpassing myself in my studies. (IMTA 1)
7. To prove to myself that I am capable of completing my college degree. (IR 1)
8. In order to obtain a more prestigious job later on. (ER 2)
9. For the pleasure I experience when I discover new things never seen before. (IMTK 2)
10. Because eventually it will enable me to enter the job market in a field that I like. (IDENT 2)
11. For the pleasure that I experience when I read interesting authors. (IMTS 2)
12. I once had good reasons for going to college; however, now I wonder whether I should continue. (AMOT 2)
13. For the pleasure that I experience while I am surpassing myself in one of my personal accomplishments. (IMTA 2)

14. Because of the fact that when I succeed in college I feel important. (IR 2)
15. Because I want to have "the good life" later on. (ER 3)
16. For the pleasure that I experience in broadening my knowledge about subjects which appeal to me. (IMTK 3)
17. Because this will help me make a better choice regarding my career orientation. (IDENT 3)
18. For the pleasure that I experience when I feel completely absorbed by what certain authors have written. (IMTS 3)
19. I can't see why I go to college and frankly, I couldn't care less. (AMOT 3)
20. For the satisfaction I feel when I am in the process of accomplishing difficult academic activities. (IMTA 3)
21. To show myself that I am an intelligent person. (IR 3)
22. In order to have a better salary later on. (ER 4)
23. Because my studies allow me to continue to learn about many things that interest me. (IMTK 4)
24. Because I believe that a few additional years of education will improve my competence as a worker. (IDENT 4)
25. For the "high" feeling that I experience while reading about various interesting subjects. (IMTS 4)
26. I don't know; I can't understand what I am doing in school. (AMOT 4)
27. Because college allows me to experience a personal satisfaction in my quest for excellence in my studies. (IMTA 4)
28. Because I want to show myself that I can succeed in my studies. (IR 4)

Appendix E: Mplus Program for SEM Model

```
TITLE:      SEM Model

DATA:       FILE IS "C:\semdata.dat";

VARIABLE:   NAMES ARE CAOS v1-v45;

    MISSING ARE ALL (99);

    USEVARIABLES ARE CAOS v1-v45;

ANALYSIS:   ESTIMATOR = MLR;

MODEL:      f1 by v1-v4;    !This line defines External Regulation factor

    f2 by v5-v10;   !This line defines Intrinsic Motivation factor

    f3 by v11-v14;  !This line defines Amotivation factor

    f4 by v15-v17;  !This line defines Identified Regulation factor

    f5 by v18-v23;  !This line defines Affect factor

    f6 by v24-v29;  !This line defines Cognitive Competence factor

    f7 by v30-v38;  !This line defines Value factor

    f8 by v39-v45;  !This line defines Difficulty factor

    CAOS on f1-f8;

    f1-f4 with f5-f8 @ 0;

OUTPUT:     STAND;
```

Appendix F: Correlations between CAOS 4 Score, Motivation and Attitude Subscales, and Demographic Variables

	CAOS	Intrinsi	Amotiv	Extern	Identifi	Affect	CogCo	Value	Difficu	Female	Class	Requir	OlderS	Colleg
CAOS 4 Score	1.000													
Intrinsic Motivation	-0.029 (.336)	1.000												
Amotivation	-0.068 (.025)	-0.049 (.103)	1.000											
External Regulation	-0.072 (.018)	0.142 (<.001)	-0.318 (<.001)	1.000										
Identified Regulation	-0.164 (<.001)	0.469 (<.001)	-0.250 (<.001)	0.451 (<.001)	1.000									
Affect	0.230 (<.001)	-0.005 (.860)	-0.122 (<.001)	-0.001 (.970)	-0.012 (.684)	1.000								
Cognitive Competence	0.265 (<.001)	-0.010 (.740)	-0.168 (<.001)	0.036 (.233)	-0.012 (.680)	0.828 (<.001)	1.000							
Value	0.187 (<.001)	0.114 (<.001)	-0.213 (<.001)	0.072 (.018)	.099 (.001)	0.444 (<.001)	0.423 (<.001)	1.000						
Difficulty	0.195 (<.001)	-0.082 (.007)	-0.068 (.024)	-0.054 (.072)	-0.079 (.009)	0.637 (<.001)	0.651 (<.001)	0.208 (<.001)	1.000					
Female	-0.213 (<.001)	0.006 (.849)	-0.158 (<.001)	0.096 (.002)	0.152 (<.001)	-0.196 (<.001)	-0.197 (<.001)	-0.091 (.003)	-0.112 (<.001)	1.000				
Class	-0.049 (.110)	-0.008 (.781)	0.111 (<.001)	-0.043 (.157)	-0.017 (.577)	-0.037 (.224)	-0.028 (.362)	-0.050 (.103)	-0.099 (.001)	-0.085 (.005)	1.000			
Required	-0.117 (<.001)	0.013 (.671)	0.024 (.436)	0.028 (.357)	0.021 (.490)	-0.150 (<.001)	-0.108 (<.001)	-0.093 (.002)	-0.126 (<.001)	0.119 (<.001)	-0.000 (.994)	1.000		
Older Student	0.001 (.974)	0.089 (.003)	-0.062 (.041)	0.030 (.321)	0.103 (<.001)	0.042 (.166)	0.063 (.037)	0.111 (<.001)	-0.062 (.041)	-0.095 (.002)	0.160 (<.001)	0.007 (.806)	1.000	
College Math & Stat	-0.092 (.002)	-0.028 (.359)	0.049 (.103)	.024 (.422)	0.037 (.216)	0.085 (.005)	0.056 (.062)	0.050 (.098)	0.035 (.248)	-0.093 (.002)	0.468 (<.001)	0.021 (.487)	0.158 (<.001)	1.000

Since these item averages will be used in a multiple regression predicting CAOS score, they are only calculated based on the n = 1094 students that completed all instruments.

* Statistics for Class variable calculated based on 1075 responses after ‘other’ responses were turned to blanks.