# UNDERSTANDING THE IMPACT OF STUDENTS' PSYCHOLOGICAL DISPOSITIONS AND BEHAVIOR ON STUDENT EXAM PERFORMANCE IN AN UNDERGRADUATE BUSINESS STATISTICS COURSE

By

Brent Marinan



A Thesis Submitted to the Faculty of the

GRADUATE INTERDISCIPLINARY PROGRAM IN STATISTICS

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

In the Graduate College

THE UNIVERSITY OF ARIZONA

2016

STATEMENT BY AUTHOR

The thesis titled UNDERSTANDING THE IMPACT OF STUDENTS' PSYCHOLOGICAL DISPOSITIONS AND BEHAVIOR ON STUDENT EXAM PERFORMANCE IN AN UNDERGRADUATE BUSINESS STATISTICS COURSE prepared by Brent Marinan has been submitted in partial fulfillment of requirements for a master's degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.



SIGNED: Brent Marinan

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

| | |
|---|---|
| | Defense date |
| Dr. Nicole B. Kersting<br>Associate Professor of Education | September 6th, 2016 |

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor Dr. Nicole Kersting for the guidance, useful comments and engagement through the learning process of this Masters thesis. Furthermore, I would like to thank the other members of my thesis committee: Dr. Joseph Watkins and Dr. Katherine Barnes for their encouragement, insightful comments, and thought-provoking questions. I am also thankful for having had professors Dr. Walter Piegorsch and Dr. Lingling An during my studies whose passion for Statistics was contagious and kept an at-risk student motivated during the more difficult moments of course content. I would also like to thank the participants of the surveys, without whose sincere responses none of this would have been possible.

Table of Contents

List of Tables

List of Figures

Abstract

Research on K-12 and college students has shown that both cognitive and non-cognitive factors affect students' academic achievement and success. This study investigates the impact of several non-cognitive factors (i.e., anxiety, beliefs, habits, motivation, context, and grit) on undergraduate students' exam performance in a public research university. Participants were 646 undergraduate students taking a statistics course designed for business students. To obtain measures of students' psychological dispositions and behaviors, students completed a survey that was developed for this study three times during the course of the semester, each time before students took one of the three course's exams. Results from the three surveys indicated that although the impact of different non-cognitive factors on exam performance decreased over the course of the semester, anxiety, beliefs and intrinsic motivation statistically significantly predicted student exam scores and the cumulative exam score. In addition, results from latent class analysis of the first wave of survey responses allowed for the determination of several diagnostic classifications of students whose academic performance, as measured by their cumulative exam scores, may benefit from early intervention.

## 1. Introduction

The transition from high school to college life can be a very challenging period in a young adult's life. Upon reaching their sophomore year, a student might expect that the steepest part of the learning curve of adjusting to academic, social, or personal transitions has passed. Unfortunately, poor adjustments to these transitions generally have negative consequences in the classroom (Aud, 2011). Academically, that steep incline continues for many of the students taking a pre-requisite 2nd business statistics class at the University's undergraduate business school. It is widely understood by students of the business school that this course is the most rigorous they take before their upper-level classes commence. Many previously successful students, as determined by academic grades, have encountered great difficulty over the class' three exams.

Just as these students face such a challenge, the business school faces its own challenges ensuring these students are adequately prepared for their upper-level business courses. In attempting to educate these students, exposure to such rigor can often times be a discouraging experience. The goal of this study is to use and build upon the considerable amount of research that has attempted to identify important non-cognitive factors to student academic success; application of these factors to undergraduate business students in explaining and predicting success, as well as early-identification of those on a negative trajectory, is explored in this study. Understanding some of the key non-cognitive constructs for student success in this class is the overarching goal of this analysis.

Mathematics scores during the undergraduate program's lower division classes are one of the strongest predictors of student retention (Lebold, 1988). Although no previous research has been done, it is understood that successful performance in the four-credit business statistics course often translates to achievement in the last two years at the business school. The class is seen by many as a rite of passage that bridges lower-level to upper-level classes, where previous students have echoed the sentiment that it has adequately prepared them, in context and rigor, for their future studies in the business school. Understanding how non-cognitive factors are linked to, or predictive of, student achievement is essential as they may suggest a diagnostic tool for intervention by instructors and educators within the business school. This, in turn, may be pertinent to contributing to the academic success and persistence in the students' challenging programs of future study.

Research has identified a considerable number of psychological factors and behaviors that can affect student achievement. Based on a review of the current literature and the specific focus of this study, the following six non-cognitive factors seemed particularly relevant for affecting students' exam performance in the undergraduate business statistics course: anxiety, specifically mathematics or statistics anxiety, motivation, both intrinsic and extrinsic, beliefs reflecting a growth or a fixed mindset, study habits, context conceptualized as social support and student employment, and grit with a focus on perseverance. The specific objectives of this study were to: 1) develop reliable measures of the non-cognitive constructs under study, 2) test various combined predictive relations among this study's non-cognitive factors and exam performance, and 3) develop a diagnostic tool to assess students' ability to succeed in the course.

Consistent with the overall goals and objectives, the study attempted to answer four research questions essential to understanding how students' disposition and behaviors affect their course performance.

1. How well do the items used in this study measure the non-cognitive factors hypothesized to affect student performance? Are the scales developed in this study, which consist of new items and existing items, internally consistent and reliable?

2. How well do the measured latent constructs explain student exam performance?

3. How stable and consistent are the explanatory results of these latent constructs, over time (i.e., different time points throughout the semester)?

4. Does the study allow identification of students who may be likely to underperform in the course? Is there an identifiable set of malleable psychological dispositions and behaviors so that an intervention can help re-orient students who might be at risk?

## 2. Literature Review

Research on the prediction of academic performance, involving more than 300 previous studies, found that a student's cognitive ability accounts for 35 to 45 percent of the variation in academic grades (Lavin, 1965). No other single indicator accounts for this much variation; however, more than half of the variation still remains unexplained (Lavin, 1965). Many subsequent studies have been directed toward identifying factors that are non-cognitive in nature to help account for the remaining variation. As a number of non-cognitive constructs have been found to be significant predictors of academic success, and it is beyond the scope of this study to capture all of such constructs, a criterion for selecting a subset was developed from more recent research. Commonly studied non-cognitive constructs include psycho-social traits, behavioral tendencies, personality traits and student affect (Cooper, 2014). It has recently been established that five general domains of non-cognitive factors emerged: academic behaviors, academic perseverance, social skills, learning strategies and academic mindsets (Nagaoka, 2013). Exploring literature from each of these domains and then examining similarities within the models preceded sketching of a conceptual map of key constructs that may most effectively relate to this course. The below section summarizes the evidence for selecting a subset of these factors that might be likely to play a role in determining student exam performance in this business statistics course.

### 2.1 Non-Cognitive Factors

**2.1.1 Anxiety.** *Anxiety* is an emotion characterized by feelings of tension, worried thoughts, and physical changes, such as increased blood pressure (APA). Research has found that anxiety about the subject of mathematics is not uncommon. Although small amounts of facilitating anxiety can help performance, anxiety is often debilitating and can have a negative effect on learning (Shipman & Shipman, 1985). Math anxiety is characterized by feelings of tension, apprehension, and fear of performing mathematics. It is also associated with delayed acquisition of core mathematics and number concepts, as well as and poor mathematics competence (Richardson and Suinn, 1972). A meta-analysis of 151 studies on grades 1-12 showed math anxiety to explain 27% of the variation in test performance (Hembree, 1990). This coefficient of determination jumped to 37% when instrument score reliabilities were near .85 (Tryon, 1980). Math anxiety is clearly an obstacle to mathematics achievement as it more than doubled the variance accounted for by a model of math fluency and experience (Ashcraft and Ridley 2005). Hence, the fact that this highly-anticipated and often-dreaded pre-business course involves mathematics and statistics should come as little surprise.

Widely accepted as a branch of mathematics, statistics is alike in producing similar characteristics of anxiety. Statistics anxiety has been defined as the feelings of anxiety encountered when taking a statistics course or doing statistical analyses involving gathering, processing, and interpreting data (Cruise, Cash, & Bolton, 1985). There are negative effects of statistics anxiety on students' learning and performance in statistics-related courses (Blalock, 1987). Students with such anxiety experience high levels of discomfort in the following situations: (1) taking notes during class lectures, (2) taking tests, and (3) doing statistical computations (Williams, 2010). A concrete example of a contributing factor to anxiety is an inability to decode terminology and symbols (Malik, 2008). Undergraduate students with high levels of anxiety may delay enrolling in statistics courses until the end of their academic programs (Onwuegbuzie, 1997). Such a delay is not possible with the 2nd statistics course, however, as it is a prerequisite for any intended business major. This study hypothesizes that students' anxiety over the subject matter will be a strong indicator of performance on the course's exams.

**2.1.2 Beliefs.** Research in recent years suggests that core *beliefs* can set up different patterns of response to challenges and setbacks (Dweck, 1999). In the present study, the first component for the belief construct was derived from the type of mindset a student has. A mindset is a set of beliefs or a way of thinking that determines one's behavior, outlook, and mental attitude (Dweck, 1988). A "growth" mindset is a belief system in which abilities, intelligence, and talents are malleable, which can be refined through effort. A person with a growth mindset believes that the brain is like a muscle that can be trained, which leads to the desire to improve (Dweck, 1999). A person with a "fixed" mindset is one who believes their talents and abilities cannot be improved through any means (Dweck, 1999). People with this mindset often avoid challenges and stick with what they know, as challenges are difficult and there is no assurance of success (Dweck, 1999). For most students, by the time they reach high school, they will have already identified the procedures that will maximize their course points with the least effort to avoid challenges (Ames, 1988).

Young students reach a higher level of mathematics achievement when they have a growth mindset. The students given techniques to develop a growth mindset "were clearly doing better than the students in the other workshop (no techniques). This one adjustment of students' beliefs seemed to unleash their brain power and inspire them to work and achieve" (Dweck, 2008, p. 215). Students who subscribe to the "entity" theory believe that intelligence is unchangeable. Those who think of intelligence as a malleable quality support an "incremental" theory (Dweck & Leggett, 1988). An experiment teaching incremental techniques to 7th graders promoted positive change in motivation and increased achievement trajectories over a two-year period, as explanation of mathematics achievement became a significant predictor ($R^2 = .45$) in path analysis (Blackwell, 2007). The expectation for this study is that stronger course performance results from those with a growth mindset tendency.

**2.1.3 Motivation.** Some studies of *motivation* in mathematics education have student goals falling into one of two categories: "performance" or "mastery" (Button, Mathieu & Zajac, 1996). Performance is an extrinsic attribute leading students to measure success externally, such as earning points, respect of peers, and external award recognition. Students are more concerned with the portrayal of their knowledge on the surface level while demonstrating their competence to others (Church, Elliot & Gable, 2001). In contrast, the category of mastery has intrinsic characteristics of student motivation, such as a desire to learn and improve at a skill or task. Here, students are interested in gaining a deeper knowledge and see academic goals as a way of furthering their own academic acumen (Cooper, 2014).

Findings have been mixed regarding which of the two categories is a better predictor of academic success. It has been demonstrated that mastery-oriented students scored lower than performance-oriented students on an introductory college psychology course exam, accounting for 7% of the total variance (Okun, 2006). The opposite finding was true, as mastery-oriented students scored higher than performance-oriented for a college business course exam, explaining a similar amount of variance (VandelWalle, 2001). Such inconsistencies may be a result of the student orientation being situational (Pintrich, 2000). For example, a final exam may orient a student to adapt more of a performance approach with the greatest concern being the overall grade, where this may not be the case for an exam earlier in the semester. Those studying motivation in learning have found that the mastery types of goals are more likely linked to perseverance as the material becomes more challenging (Middleton & Toluk, 1999). Because this pre-requisite business course rapidly builds on abstract fundamental concepts throughout the semester and challenges students with exams of increasing difficulty, this study hypothesizes that students having high intrinsic motivation and low extrinsic motivation would result in better exam performance.

**2.1.4 Habits.** *Habits* are mainly external factors that facilitate the study process, such as sound study routines, including how often a student engages in study sessions, self-evaluates, and reviews the class material (Cerna, 2015; Credé, 2008). A student's habits are a significant variable determining their academic performance (Baquiran, 2011). Sound and persistent habits help improve a student's performance, reduce exam anxiety, enhance ability, and develop confidence in herself (Reed, 1996). Arriving at college, many students assume that previous habits leading to academic success in high school will readily translate to a college setting. However, the realization that college is very different soon begins to set in for many students (Robinson, 2010). This is the case for many pre-business students enrolled in this 2$^{nd}$ business statistics course. As the course's exams are multiple choice, some business students might be inclined to take what Nonis refers to as a "surface" approach (r = -.21 on GPA), as opposed to a "deep" approach (r = .38), for the first exam (Nonis, 2012).

Habits showed the most consistent relationships with undergraduate academic performance (Entwistle, 1971). The amount of time invested in successful habits is of great importance (Nausheen, 2002). Weekly study time has decreased almost 40% over the past 50 years (Babcock, 2010). With such a decrease in study time, not postponing what is necessary to reach a goal (procrastination) grows in importance. Procrastination is a self-handicapping behavior that leads to wasted time, increased anxiety, and poor academic performance (Ozer, 2011). With over 50% of undergraduate students admitting to consistent and problematic procrastination, their tendencies are a barrier to academic success (Day, 2000). Given these findings, it is hypothesized that habits, including procrastination, will affect student performance.

**2.1.5 Grit.** A more recent contribution in research to academic achievement is the non-cognitive indicator *grit*. It builds upon the personality factor conscientiousness (a Big Five personality trait) in its predictive ability of academic success (Conard, 2006). "Grit shares the achievement aspect of conscientiousness, but grit requires sustained effort and interest in goals, notwithstanding failure, lack of progress and feedback, and difficulty" (Chang, 2014, p. 22). It is defined as "trait-level perseverance and passion for long term goals" (Duckworth, 2007, p. 1087). Grit has been linked to success with Ivy League undergraduates (r = .25), West Point cadets (r = .19), and National Spelling Bee participants (Duckworth, 2007). It has also been suggested as a stronger indicator for long-term student performance than conscientiousness or intelligence (Duckworth, 2007). Two factors, "consistency of interests" and "perseverance of efforts", make up the composition of the Grit Scale. As our research was less concerned about longer-term achievement, which was reflected more in the former subscale, our focus was mainly on the latter factor of the Grit scale. It is hypothesized that, even though previous studies of grit focused more on the stronger-performing students, this indicator will be a positive indicator of student performance in this study's population.

**2.1.6 Context.** It would be very difficult to ignore the effect on academic performance that the home, work and school environment of a student has. Until recently, the *contextual* factor, "social support", has generally been neglected in educational models or measured as general perceived support and general perceived social inclusion (Tinto,2004). In this generation of technology, its importance as a buffer against stress for emerging adults is growing (Chao, 2012). A meta-analysis study measuring social activity and social connection indicated a strong relation ($r = .26$ and $r = .22$) to academic performance (Robbins, 2006). Another study showed that support variables had no effect on years to complete the diploma, though it used very general measures of social networks and social support (Eggens, 2008).

A 2nd contextual factor, student employment, has shown a relationship with academic performance. Participation in the workforce demonstrated a positive relationship with GPA ($r = .12$) (Cannabal, 1998). A similar study found that students who do not work earned lower grades than students who work part-time (Stern, 1991). Student grades tended to improve as students worked more hours per week, up to a total of 20 hours per week (Van de Water, 1996). Additionally, students who worked 10 to 20 hours per week performed better academically than students who worked less than 10 hours, more than 20 hours, or not at all (Van de Water, 1996). Thus, because these studies have shown that social support and employment contribute positively to academic performance, it is hypothesized that both of these contextual factors will enhance exam performance for this study's students.

## 3. Method

### 3.1 Background

Participants included 646 students (59% male, 41% female) undergraduate pre-business students. Most of the students are in their 2nd year of school and fall within 19 to 20 years in age. Overall in the business school, 90% of the students are enrolled full-time, 51% are state residents, and 60% are White (19%, Hispanic/Latino) (Quick Summary, 2014). These three demographics can be seen to approximate this study's participants.

The entirety of the data was collected during the spring semester of this business statistics course. The course is comprised of three different sections. Two of them are taught by the same professor, and the third section is taught by a different professor. All three of the sections are included in the study. To take this 2nd course in business statistics, students had to pass two pre-requisite mathematics courses and one statistics course (Applied College Algebra, Calculus Concepts & Statistical Inference). Participants' overall grade in these three pre-requisite courses, as well as their last high school mathematics course, is as follows: approximately 43% achieved a level of "A", 49% achieved a level of "B", and 8% achieved a level of "C". The mean grade for all participants in the four courses approximated a B+. During the course of the semester, over two-thirds of the students held a part-time job, with approximately half of them working less than

10 hours per week and the other half working at least 10 hours per week. With the majority balancing school and work, almost 90% of the students agree, on some level, that they had a strong support system in place during the semester. Lastly, historically, just over two-thirds of the students taking this course will be admitted to the business school.

### 3.2 Procedure

Data for the six non-cognitive constructs were gathered from the participants through the use of Turning Point clickers, during three different periods, over the course of the semester. For all three student sections, a survey was distributed one week prior to each exam. The questionnaires were administered during the class lab times. Time constraints during the class time limited the amount of questions to be asked. The response rate was 94%. Students were notified that their attendance for the day would be captured through participation in this survey. Each assessment took approximately 10 minutes to complete. After discussion with an Institutional Review Board (IRB) manager, it was determined that an application for review would not be needed for this study. In the compilation of data, only the students' clicker identifier was used. Involved student names were not identified, and the clicker data was exported to Excel for analyses.

Several preliminary steps were taken to maximize data quality. The data were first evaluated for missing item responses or duplicate survey entries. Participants who failed to respond to more than 25% of the items or who completed any survey more than once, ($n = 39$) were omitted (Cooper, 2014). The number of participants then totaled 646. Once the study's analysis began, the number of observations in a given analysis varied as a result of STATA's Maximum Likelihood (ML) (list-wise deletion) selection method. The percentage of missing data using this method varied by model, ranging from 10.5% to 60.3%, which is considerable and might impact the results and findings of this study. This is discussed further in the Limitations section.

### 3.3 Measures

**3.3.1 Student Course Exams**. Scores on the semester's three exams were used as the measures of achievement. Each exam covers different content. There is no cumulative exam. Each exam included 25 multiple-choice conceptual questions (2 points per question) and 10 multiple-choice word problems (4 points per problem, involving calculations), collectively worth 90 points. The questions on the exams were identical for all students. To understand whether data from the three sections could be combined for analysis, a one-way ANOVA test was performed. The test showed no significant differences between the population means (See Appendix A). This was also the case for the three sections' cumulative exam scores (See Appendix A).

The above finding was used to allow the data from all three sections to be pooled. Once pooled, each period of exam scores and cumulative exam scores were standardized accordingly with their respective normal distributions. The remaining components of the course (i.e., homework, cases, and attendance) were not considered for inclusion as nearly 80% of this study's students have cumulative perfect scores in those categories, which suggests low variability. Patterns of the study's non-cognitive variables may not manifest themselves if such measures of achievement are conflated.

**3.3.2 Student Survey of Social Psychological Attributes and Behavior.** A set of scales designed to measure relevant non-cognitive variables was developed. Based on theoretical considerations, the first survey consisted of a 24-item questionnaire (See Appendix B) that attempted to measure five individual constructs: *anxiety*, *beliefs*, *motivation*, *habits*, and *context*. Restructuring the 1st survey led to re-classification of *motivation* into the categories *intrinsic* and *extrinsic,* as well as adding the factor *grit* to the 2nd (37 items, See Appendix C) and 3rd surveys (39 items, See Appendix D). This resulted in the 2nd and 3rd surveys comprising seven latent constructs. The survey items used a six-point Likert scale (Dweck, 2007), ranging from 1 (strongly disagree) to 6 (strongly agree). In order to simplify the interpretation of the relationship between the variables as well as to prepare variables to be combined into a scale, reverse coding was used for all negatively-scored items.

As expected, not all of the items initially included in the survey worked sufficiently well as indicators of their respective latent constructs. The seven latent constructs and the final number of indicators in each construct follows: *anxiety* (3), *beliefs* (2), *context* (2), *extrinsic motivation* (3), *intrinsic motivation* (2), *habits* (3), and *grit* (3). The exact wording of the items is shown in Table 1. The numbering of the indicators, as seen in Table 1, reflects the indicators' numbering sequence as they originally appeared in the study's surveys. For example, although anxiety had six items in the survey, only questions 1, 4 and 5 were retained as indicators of anxiety and used for analyzing this construct's effect on student performance.

Most items included in the survey were taken from relevant existing instruments. However, the surveys also included items based on the author's intuition over several years of lab instruction in the business statistics course. Data for the indicators was gathered from the participants using the measures outlined below.

TABLE 1
*List of Selected Items for Non-Cognitive Constructs*

| Factor | Item |
|---|---|
| Anxiety | 1: Mathematics makes me feel uncomfortable and nervous. |
| | 4: My performance tends to suffer when I have to race against deadlines. |
| | 5: The new symbols and Greek letters, in this class, can make me very anxious. |
| Beliefs | 1: You have a certain amount of intelligence and you really cannot do much to change it. |
| | 2: You can always greatly change how intelligent you are. |
| Context | 1: I feel like I have a strong social support network in University, and surrounding city, to help me through any difficult times. |
| | 3: In recent weeks, I have been working at a job ______________ per week. |
| Ext Mot | 1: I am doing pre-business (vs. other majors) to have the best opportunities to find a job upon graduation. |
| | 2: I just want to avoid doing poorly, grade-wise, in this class. |
| | 3: I am very concerned with my parent's thoughts of my grades. |
| Int Mot | 1: I would take a Statistics course in future as an elective (non-required) course. |
| | 2: I am more concerned with learning the content of the class than earning a good grade. |
| Habits | 2: I have been regularly doing non-graded practice problems for each chapter. |
| | 3: I have sufficiently timed myself when working out problems in preparation for the exam. |
| | 4: I have been consistently reading the required chapters in preparation for the exam. |
| Grit | 3: I am a hard worker. |
| | 5: I am diligent. |
| | 6: I finish whatever I begin. |

**3.3.2.1 *Anxiety.*** Item 1 came from the 10-item Mathematics Anxiety Scale (Betz, 1978), which was found to have acceptable internal consistency and test-retest reliability (Dew, Galassi & Galassi,1984; Pajares & Urdan, 1996). Item 4 was taken from the 16-item New Active Procrastination Scale. The level of reliability (Cronbach's α: discussed in "Models" section) assessing the four dimensions ranged between .70 and .83, providing evidence of acceptable internal consistency (Choi J., 2009). It also exhibited an acceptable reliability coefficient of .80 (Choi J., 2009). Item 5 came from an interview assessing statistical anxiety of a student (Malik, 2008) no measure of reliability was reported. All three of the items in this construct were reverse coded, meaning that a high score on anxiety relates to a low amount of anxiety.

**3.3.2.2 *Beliefs*.** Items 1 and 2 were drawn from the Theory of Intelligence scale (Dweck, 1999). Item 1 was taken from the *incremental theory* category and reverse coded, while Item 2 was taken from the *entity theory* subscale. The reliability of the scales was estimated at .78 and .77, respectively (Dweck, 1999).

**3.3.2.3** ***Context.*** All items were created by the study's author. The three items used in this study's survey were developed on previously-mentioned theory regarding social support and student employment.

**3.3.2.4** ***Habits***. Items 2, 3 & 4 of this construct were based on author intuition regarding past student success in the course.

**3.3.2.5** ***Motivation***. The 1$^{st}$ survey originally only had one construct for motivation. After the 1$^{st}$ survey, exploratory factor analysis suggested that items measured two subscales, *intrinsic* and *extrinsic* motivation. Item 2 of the extrinsic subscale came from the 16-item Achievement Goal Orientations Questionnaire, which measured an alpha of .79 (Cao, 2012). The study's author created the other four items chosen for the two subscales' analysis. For the analysis, two items from *intrinsic motivation* and three items from *extrinsic motivation* were considered. All the items from *extrinsic motivation* were reverse coded, meaning a high score on these items reflects a low amount of extrinsic motivation.

**3.3.2.6** ***Grit***. This construct was added to the 2$^{nd}$ and 3$^{rd}$ surveys and did not appear in the first survey. Items 3, 5, & 6 were taken from the "Perseverance of Effort" subscale of the 12-item Grit Scale (Duckworth, 2007). In that study, the Grit Scale exhibited a high internal consistency (. 85).

Items drawn from existing instruments have demonstrated adequate internal consistency and reliability in their respective studies. The reliabilities observed in the present data will be discussed in the Results section. In addition to the non-cognitive items, one item included in the 1$^{st}$ survey measured prior achievement. The item, *prev*, concerned the students' overall grade in their four most recent mathematics/statistics-related courses. This item only appeared on the 1$^{st}$ survey and ranged on a scale of 1 to 8, with a grade of "C-" having a value of 1, and a grade of "A" having a value of 8. Details on the statistical procedures to determine the final set of indicators for each construct follows below in the Statistical Models and Approaches section.

## 3.4 Statistical Models and Approaches

**3.4.1 Factor Analysis.** Although the items selected for this study came from scales that were found to produce reliable scores, extraction of individual items across those studies and compilation in this study required a reassessment of the measurement error. To understand how well the surveys' items measured the non-cognitive constructs used in this study, the underlying factor structure of the survey items was the first criterion to consider. Using the items of the respective surveys in order to derive the final set of indicators used in the study's analysis, a factor analysis was conducted for each measured latent construct.

Factor analysis is a method of data reduction, which removes redundancy from a set of correlated variables. This analysis identifies underlying latent factors, relatively independent of each other, that are reflected through the observed variables (Rummel, 1967). It is based on the correlation matrix of the variables involved and requires a large sample size before stability occurs, where a minimum of ten observations per variable is a good rule of thumb. (Fidell & Tabachnick, 2001). For this study, 250+ is considered fair to good (Comrey & Lee, 1992) when attempting to fit various 1-factor models.

There are several different methods that can be used to uncover a pattern through factor analysis (Tabachnick & Fidell, 2001). This study first used factor analysis in an exploratory manner for each latent construct to understand how well the items assumed to measure a given construct clustered together. One by one, items that constituted a single-item factor or had high cross-loadings on multiple factors were removed. In constructing a latent variable, high communalities (i.e., factor loadings without cross loadings) are desirable. Though unlikely to occur in real data, item factor loadings are considered "high" if they are all .8 or greater (Ava & Velicer, 1998). In the social sciences, more common favorable magnitudes are low to moderate loadings of .40 to .70 (Costello & Osborne, 2005). If an item has a loading of less than .40, the researcher should consider why that item was included in the data and decide whether to drop it or add similar items for future research (Costello & Osborne, 2005). One study cites .32 as a good rule of thumb for the minimum loading of an item, which equates to approximately 10% overlapping variance with the other items in that factor (Tabachnick & Fidell, 2001).

After completing factor analysis for exploratory purposes, each set of items determined to be reasonable indicators of respective non-cognitive constructs were factor-analyzed in a confirmatory procedure. This involved fitting a 1-factor model separately for each construct. Figure 1 illustrates a confirmatory 1-factor model, *Factor_1,* where the arrow represents the latent construct's effects on each of the three indicators. The factor loading is the weight of each effect by the latent construct onto the respective indicator. The factor loading is essentially the relationship of each observed variable to the underlying factor. For this study, a one factor model was used for each respective construct, and quality of fit will be discussed in the Results section.

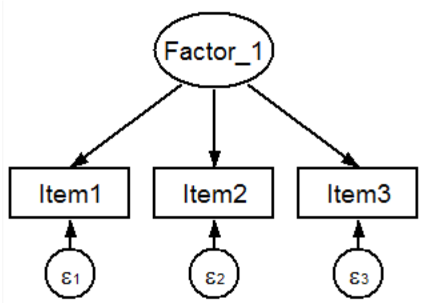

*Figure 1 Single Factor Model (3 indicators, error terms)*

In exploring the data for patterns, factor-analyzing involving one factor can be represented algebraically. If the number of observed variables, X, are dependent on a latent variable, F, where λ represents the loading for each $X_j$ and the measurement error for $X_j$ is represented by $e_j$, then the series of equations can be expressed as follows:

$X_1 = \lambda_1 F + e_1$
$X_2 = \lambda_2 F + e_2$
….
$X_m = \lambda_m F + e_m$

Under this series, factor analysis assumes that (a) the measurement error has constant variance, (b) there is no association between the factor and the measurement error, and (c) the observed variables are independent of each other (Rummel, 1970). The goal of factor analysis is to ultimately find latent variables common to particular sets of observed variables and the effect that the respective latent construct has on its observed variables. In addition to determining the loadings, factor analysis also generates factor scores for each respondent, which indicate a respondent's placement or standing on the latent construct. The above factor-analyzing procedures were performed, as well, for items from the 2nd and 3rd surveys to understand the stability of the measures over time.

Another measure of how well the observable items measure the same latent construct is termed "internal consistency". This measurement of reliability is often measured with Cronbach's alpha, a statistic calculated from the pairwise correlations between items (Cortina, 1993). Cronbach's alpha provides information similar to factor analysis, although in the form of a single value. Theoretically, alpha may be expressed as a function of the parameters of the hierarchical factor analysis model. This allows for a general factor that is common to all of the items of a measure, in addition to group factors that are common to some but not all of the items of a measure (Li, Revelle, Yovel & Zinbarg, 2005).

Cronbach's alpha is known as an internal consistency estimate of reliability of item scores because it will generally increase as the intercorrelations among items increase (Cortina, 1993). Because intercorrelations among items are maximized when all items measure the same construct (Cortina, 1993), Cronbach's alpha can be seen as a measure to indicate the degree to which a set of items measures a single unidimensional latent factor. Alpha is most appropriately used when the items measure different substantive areas within a single factor (Cortina, 1993) and will be applied as such in this study. For example, in modeling the factor *anxiety*, the different substantive areas in this study were 1) performance anxiety upon meeting a deadline, 2) overall mathematics anxiety and 3) anxiety due to new and unfamiliar symbols used in mathematics.

**3.4.2 Structural Equation Modeling.** To study the effects of non-cognitive factors on student performance, structural equation modeling (SEM) was employed. SEM is an extension of multiple regression in which the relationships between multiple factors are evaluated simultaneously. SEM, using latent variables to account for measurement error, has been used more recently in research to examine the relationships between multiple factors in education. In most cases, such models have served as a means to test existing theoretical models of the relationships between constructs of different types (i.e., goals, study strategies, and academic achievement) (Dupeyart & Marine, 2005). When increasing the number of latent variables in a model, a significant increase in the minimum sample size is needed, even when moving from one to two factors (Wolf, 2013). Sample size requirements ranging from 30 (Simple confirmatory factor analysis with four indicators and loadings around .80) up to 450 cases (mediation models) were considered sufficient. (Wolf, 2013). The maximum likelihood (ML) method was used for estimating the models. This method, in STATA[1] (StataCorp LP, 2016), allows for relaxation of the assumption of joint normality for all variables. Because of ML's list-wise deletion method, the number of observations varied by model.

As multiple factors are evaluated simultaneously in SEM, the path diagram in Figure 2 represents a basic two-factor model in which each factor represents the exogenous variable. The arrow directed from each exogenous variable to the single endogenous variable represents each factor's direct effect on the endogenous variable. The curved two-headed arrow from each latent variable indicates some possible association between the two exogenous variables.

[1] With the exception of Latent Class Analysis, all statistical analyses was performed using STATA.

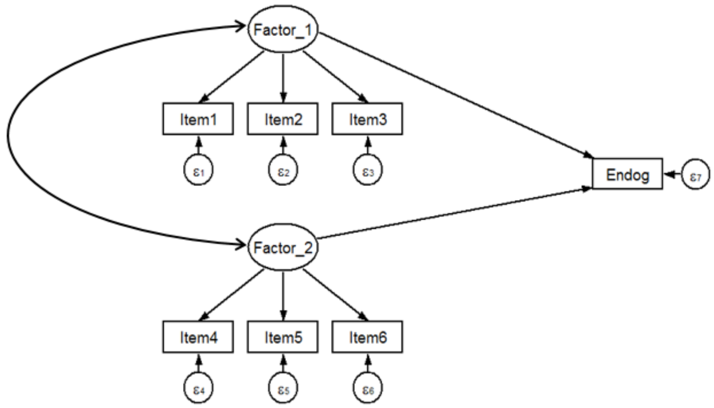


*Figure 2 Two Factor SEM with Endogenous Variable*

**3.4.3 Latent Class Analysis.** To identify students who might be at risk of performing poorly in the course, latent class analysis (LCA) was used. Unlike factor analysis, which takes a "variable-centered" approach to identify associations among variables, LCA identifies similar response patterns to a series of data points and attempts to classify individuals into homogeneous groups (Fosnacht, 2013). More practically, LCA uses a person-centered approach to identify groups of individuals who share common attributes (Hoff & Laursen, 2006). LCA also attempts to produce fitted probabilities of class membership for each individual.

The parameters in a latent class model consist of unconditional and conditional probabilities. The conditional probabilities comprise the measurement portion of the model which characterize the distribution among the indicators conditional on the latent classes (Magidson & Vermunt, 2004). The unconditional probabilities help describe the distribution of the latent variables. In order to improve the description of the latent variables, a multinomial logit model is used to express these probabilities as a function of one or more exogenous variables, Z, called covariates (Dayton and McReady, 1988). Addition of covariate(s) leads to a higher proportion of individuals assigned to the correct class, even if class separation is poor, and can help to correctly identify the number of classes (Lubke and Muthén,_2007). Description of latent variable distributions can also be improved with the addition of more indicators leading to "more converged and proper replications, as well as less parameter bias" (Geiser & Wurpts, 2014). In summary, the use of a larger number of high-quality indicators and the inclusion of at least one strong covariate positively affect LCA model estimation (Wurpts & Geiser, 2014). A basic LCA model is shown below (Figure 3).

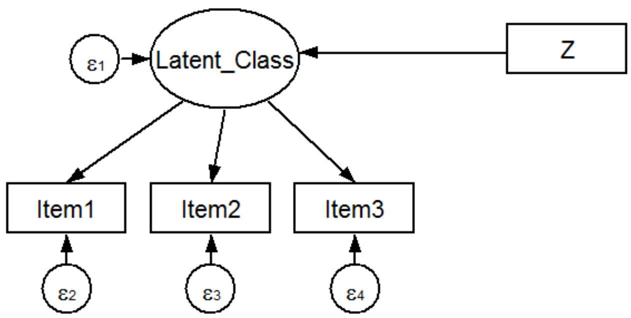

*Figure 3 Latent Class Model with Covariate, Z*

Although the literature suggests LCA is not as frequently implemented in education research as SEM modeling, LCA seems well-suited for examining relationships between non-cognitive factors and student performance. For the LCA, this study relied on LatentGOLD 5.1 software (Statistical Innovations, 2016), which uses maximum likelihood estimation to fit a hypothesized model in which membership within a specified number of latent classes is related to performance on certain measures.

The goal of LCA is to determine the smallest number of latent classes, T, that is sufficient to account for the relationships among the observed variables. It typically involves minimizing the within-cluster variation and/or maximizing the between-cluster variation (Magidson & Vermunt, 2002). The analysis typically begins by fitting the T = 1 class baseline model ($H_0$), which specifies mutual independence among the variables (Magidson & Vermunt, 2004).

Model $H_o$ : $\pi_{ijkl} = \pi_i^A \pi_j^B \pi_k^C \pi_l^D$

Assuming that this null model does not provide an adequate fit to the data, a 1-dimensional model with T=2 classes is then fitted to the data. This process continues by fitting successive models to the data, each time adding another dimension by incrementing the number of classes by one, until the simplest model found provides an adequate fit (Magidson & Vermunt, 2004). Several complementary approaches are available for assessing the fit of latent class models. The most widely used approach utilizes the likelihood ratio chi-squared statistic, $L^2$, to assess the extent to which maximum likelihood estimates for the expected cell frequencies, $F_{ijkl}$ , differ from the corresponding observed frequencies (Magidson & Vermunt, 2004), $f_{ijkl}$:

$$L^2 = 2 \sum_{ijkl} f_{ijkl} \ln(f_{ijkl} / F_{ijkl})$$

A model fits the data if the value of $L^2$ is sufficiently low to be attributable to chance (within normal statistical error limits; generally, the .05 level) (Magidson & Vermunt, 2004).

# 4. Results

## 4.1 Summary Statistics

Descriptive statistics for all of this study's variables are summarized below in Tables 2, 3 and 4. Table 2 shows the means and standard deviations of the non-cognitive constructs, the cognitive variable, *prev,* and the performance measures. Table 3 illustrates the means and standard deviations of each construct's indicators, over the course of three waves. Finally, Table 4 displays the bivariate correlations of the study's variables during wave 1.

TABLE 2
*Summary Statistics for Non-Cognitive & Cognitive (prev) Variables, and Exams*

| | *Wave1* | | | *Wave2* | | | *Wave 3* | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | n | Mean | SD | n | Mean | SD | n |
| Anxiety | 3.40 | 1.40 | 535 | 3.50 | 1.43 | 543 | 3.53 | 1.37 | 473 |
| Beliefs | 4.25 | 1.30 | 546 | 4.20 | 1.25 | 543 | 3.80 | 1.20 | 482 |
| Context | 4.35 | 1.55 | 534 | 4.20 | 1.50 | 540 | 4.45 | 1.50 | 449 |
| Ext Motiv | 2.00 | 1.13 | 544 | 2.17 | 1.13 | 538 | 2.10 | 1.10 | 443 |
| Habits | 3.13 | 1.47 | 543 | 3.27 | 1.50 | 579 | 3.47 | 1.50 | 441 |
| Int Motiv | 3.05 | 1.40 | 567 | 2.90 | 1.50 | 543 | 3.15 | 1.45 | 473 |
| Grit | --- | --- | | 4.93 | 0.90 | 543 | 5.07 | 0.83 | 435 |
| Prev | 5.85 | 1.56 | 541 | --- | --- | | --- | --- | |

*Note*. *Grit* was not measured during wave 1. *Prev* only measured in wave 1.

As displayed in Table 2, a mean of 5.85 in the previous mathematics/statistics coursework (*prev*) by participants in this study approximates a B+ average. In the business school's grading scale, a B+ average translates to a range between 87% and 90%. Such measured success in *prev* might suggest that students entering this course would have a low amount of anxiety. Table 4 (below), however, shows a moderate correlation (.37) between *prev* and item 1 of the *anxiety* construct ("Mathematics makes me feel uncomfortable and nervous"). Additionally, wave 1's mean for the overall *anxiety* construct (3.4) reveals a moderate amount of anxiety. *Anxiety's* mean slightly increased (improved) over the subsequent waves, while the mean scores for the 1st two exams (66.84% & 64.19%) were approximately twenty points lower than that of their *prev* course grades. This may suggest that once students took the 1st exam, they felt more comfortable regarding their expectations for the remaining exams. Inspection of *anxiety's* indicators from Table 3 reveal item 1 (paraphrase item 1) has the highest value (lowest amount of anxiety) over the course of the semester. Item 3 (anxiety over new symbols and Greek letters) had the biggest improvement from wave 1 to wave 3 (3.4 to 3.8). This would suggest that students' anxiety lessened as they were increasingly exposed to new symbols. As seen in tables 2 and 3, the number of observations in each construct's period changed and could help to explain some of the changing constructs' means. Overall, wave 3 had the fewest amount of observations.

TABLE 3
*Mean and SD of Non-Cognitive Items (over 3 waves)*

| | Wave 1 | | | Wave 2 | | | Wave 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| Item | Mean | SD | n | Mean | SD | n | Mean | SD | n |
| A1 | 4.1 | 1.4 | 578 | 3.8 | 1.4 | 543 | 3.9 | 1.4 | 474 |
| A4 | 2.7 | 1.3 | 535 | 2.9 | 1.4 | 585 | 2.9 | 1.3 | 473 |
| A5 | 3.4 | 1.5 | 540 | 3.8 | 1.5 | 585 | 3.8 | 1.4 | 474 |
| B1 | 4.2 | 1.3 | 571 | 3.8 | 1.4 | 543 | 3.3 | 1.3 | 491 |
| B2 | 4.3 | 1.3 | 546 | 4.6 | 1.1 | 563 | 4.3 | 1.1 | 482 |
| C1 | 4.9 | 1.1 | 570 | 4.8 | 1.1 | 540 | 5.0 | 1.0 | 482 |
| C3 | 3.8 | 2.0 | 534 | 3.6 | 1.9 | 544 | 3.9 | 2.0 | 449 |
| EM1 | 1.8 | 1.0 | 566 | 2.1 | 1.1 | 538 | 2.0 | 1.0 | 445 |
| EM2 | 1.7 | 1.0 | 544 | 1.8 | 0.9 | 624 | 1.9 | 1.1 | 484 |
| EM3 | 2.5 | 1.4 | 547 | 2.6 | 1.4 | 580 | 2.4 | 1.2 | 443 |
| H2 | 2.5 | 1.4 | 543 | 3.3 | 1.5 | 579 | 3.4 | 1.5 | 441 |
| H3 | 3.4 | 1.5 | 560 | 2.9 | 1.5 | 583 | 3.3 | 1.5 | 443 |
| H4 | 3.5 | 1.5 | 559 | 3.6 | 1.5 | 583 | 3.7 | 1.5 | 437 |
| IM1 | 2.8 | 1.4 | 567 | 2.6 | 1.6 | 581 | 3.0 | 1.5 | 474 |
| IM2 | 3.3 | 1.4 | 537 | 3.2 | 1.4 | 543 | 3.3 | 1.4 | 473 |
| G3 | -- | -- | | 5.1 | 0.9 | 543 | 5.2 | 0.8 | 435 |
| G5 | -- | -- | | 4.9 | 0.9 | 576 | 4.9 | 0.8 | 446 |
| G6 | -- | -- | | 4.8 | 0.9 | 587 | 5.1 | 0.9 | 444 |

*Note.*
A= Anxiety
B = Beliefs
C = Context
IM = Intrinsic Motivation
EM = Extrinsic Motivation
H = Habits
G = Grit

Investigating the *beliefs* construct (Table 2), a mean of 4.25 reveals a favorable level of student growth mindset in wave 1 (before 1$^{st}$ exam). The mean value for beliefs, however, decreased (worsened) over the next two waves with the biggest drop occurring during wave 3. This may be attributed to the 2$^{nd}$ exam score being lower than the (already low) 1$^{st}$ exam score. Some students' outlook and mental attitude may have decreased when they were unable to see their efforts produce certain results by the 2$^{nd}$ exam. For highly extrinsically motivated students the sentiment that improvement, in terms of performance measurement, has not occurred may push some of them towards a fixed mindset. A closer inspection reveals a contradiction between the *beliefs* indicators' means. Specifically, the mean for item 1, measuring the level of students' fixed mindset, decreased substantially after both the 1st and 2nd exams, while the mean for item 2, representing a growth mindset, increased after the 1st exam. These trends appear contradictory since all the necessary item(s) have been reverse coded, i.e. item 1. While the mean was decreasing over time, the *beliefs'* standard deviation was the only construct to continually decrease over all the waves. Further evaluation of this construct may be warranted in the measurement section of this study.

The *context* construct of this study had relatively high scores over the course of the semester, with the highest coming in the last wave (4.5). Students generally felt they had a solid support system (item 1), with the majority working at a part-time job (item 3) not more than ten hours per week. Intuitively, item 3's standard deviation being twice that of item 1 is no surprise given the wide range of student employment hours, or no employment at all.

As previously mentioned, the students in this study have very high *extrinsic motivation* as shown with the lowest values of all the constructs (Table 2). With reverse-coding, a low value in this construct translates to a high amount of extrinsic motivation. This trend matches intuition that as business school candidates, they would rank very high on external motivation as it pertains to grades as well as career choices. Even so, the students' extrinsic motivation did become lower (mean increased) overall during the course of the semester, which theories suggest should contribute to improved academic performance.

The *habits* construct's beginning mean of 3.13 (wave 1) may not provide as much substantive information as that of subsequent waves. After the 1st exam, *habits*' mean increased to 3.27 (wave 2) and further increased to 3.47 in wave 3. This might be best explained by the exam 1 and exam 2 scores, which were very low relative to their *prev* grades in coursework. Though, the correlation between habits (wave 2) and exam 1 were very low (not shown in tables), low exam scores may have indiscriminately prompted an improvement in *habits*. A closer inspection of *habits'* indicators reveals the biggest contribution in its improvement came from item 2 (doing non-graded practice problems). Item 2 jumped from a mean of 2.5 in wave 1, to 3.3 and 3.4 in waves 2 and 3. This supports speculation that students increasingly realize they must implement a deep-level approach (Nonis, 2012) to studying in order to succeed.

The construct *grit,* added to the 2nd and 3rd waves, had the highest means of all the constructs while also having the smallest standard deviations. This means that the students in this study have a very high level of perseverance. This might explain the tendency for skewness in its distribution, where the high means may result from some bias in the students' perceptions of their level of grit. The low amount of variation might also suggest grit may not be a likely predictor of exam performance.

As it concerns the standard deviations of the items from the 1st wave to the 2nd wave, 1) six items remained at the same amount, 2) three of the items had standard deviations that decreased in value, and 3) five items experienced an increase in standard deviation. This finding was similar to changes from the 2nd to the 3rd waves. Although, overall, the standard deviations of this study's constructs remained relatively constant over the course of the semester, the standard deviations of the exam showed a consistent increase where the 3rd exam's standard deviation (15.78) was 12.9% higher than that of the 1st exam (13.98). This increase in dispersion may suggest that as the semester progressed, students that continued not grasping fundamental concepts may not have had the necessary foundation for building the tools to be successful in the course. This would suggest a result of declining grades over the course of the semester.

From the bivariate correlations (of wave 1 indicators) summarized in Table 4, the singular cognitive item, *prev* (previous course grades), had a moderate relationship with the course's exam grades, as expected. This item also had a moderate association with item 1 of the *anxiety* construct. Further, it is evident that the performance outcomes are associated with some of the non-cognitive variables of interest in this study, most notably in the construct of *anxiety*. Items in *habits* and *intrinsic motivation* also seem to suggest an adequate amount of correlation. Some notable amount of negative correlation can be seen between the indicators of *beliefs* and *extrinsic motivation*. This is noteworthy because such negative correlation should be relatively non-existent among indicators after consideration of reverse coding. This negative correlation implies the association that as a mindset is moving from growth to fixed, extrinsic motivation is lessening. Overall, the remaining intercorrelations among constructs do not appear to warrant any attention as they do not meet even low classifications of correlation effects.

The constructs *anxiety*, *intrinsic motivation*, *habits* and *beliefs* showed relatively stronger associations with *exam1*, than did *context* or *extrinsic motivation*. The relative strength of these associations between constructs and performance were consistent through the remaining two waves, though, beliefs dropped off in its strength. Additionally, the introduction of *grit* in the 2nd wave showed a relatively strong association with respective performance measures. Because of the above, the factors having stronger correlations with the respective exams than others, were considered for the study's predictive models.

TABLE 4

*Descriptive Bivariate Summary for Main Study Variables (Wave 1)*

| | Prev | A1 | A4 | A5 | B1 | B2 | C1 | C3 | EM1 | EM2 | EM3 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Prev | 1.00 | 0.00 | | | | | | | | | |
| A1 | 0.37 | 1.00 | | | | | | | | | |
| A4 | 0.10 | 0.00 | 1.00 | | | | | | | | |
| A5 | 0.07 | 0.21 | 0.29 | 1.00 | | | | | | | |
| B1 | -0.06 | -0.01 | 0.01 | -0.07 | 1.00 | | | | | | |
| B2 | -0.12 | -0.06 | -0.04 | -0.08 | 0.47 | 1.00 | | | | | |
| C1 | -0.15 | -0.06 | -0.07 | -0.11 | 0.11 | 0.17 | 1.00 | | | | |
| C3 | 0.09 | 0.08 | 0.03 | 0.00 | -0.11 | -0.09 | -0.11 | 1.00 | | | |
| EM1 | 0.07 | 0.21 | 0.31 | 0.10 | -0.03 | -0.16 | -0.09 | -0.09 | 1.00 | | |
| EM2 | 0.03 | 0.11 | -0.08 | 0.09 | -0.13 | -0.21 | -0.03 | 0.14 | 0.05 | 1.00 | |
| EM3 | -0.02 | 0.22 | 0.04 | -0.08 | -0.03 | -0.15 | -0.02 | -0.24 | 0.17 | 0.21 | 1.00 |
| H2 | 0.14 | -0.14 | -0.11 | -0.18 | 0.00 | 0.13 | 0.05 | 0.02 | -0.12 | 0.00 | -0.16 |
| H3 | 0.11 | -0.04 | -0.12 | 0.01 | 0.02 | -0.05 | 0.01 | 0.08 | -0.17 | 0.19 | -0.12 |
| H4 | 0.06 | 0.10 | -0.14 | -0.21 | 0.04 | 0.11 | -0.03 | -0.12 | -0.03 | 0.00 | 0.02 |
| IM1 | 0.19 | 0.15 | 0.02 | 0.18 | -0.02 | -0.10 | -0.02 | -0.04 | 0.12 | 0.05 | 0.05 |
| IM2 | -0.01 | -0.04 | -0.11 | -0.02 | -0.14 | 0.04 | 0.00 | 0.10 | -0.04 | 0.04 | -0.18 |
| Exam1 | 0.40 | 0.37 | 0.18 | 0.25 | -0.08 | -0.18 | -0.12 | 0.08 | 0.07 | 0.07 | 0.11 |
| Exam2 | 0.38 | 0.32 | 0.08 | 0.14 | -0.05 | -0.05 | -0.11 | 0.04 | -0.04 | 0.04 | 0.13 |
| Exam3 | 0.34 | 0.31 | -0.07 | 0.03 | 0.09 | -0.03 | -0.12 | 0.22 | -0.09 | -0.02 | 0.03 |
| Cumul | 0.46 | 0.41 | 0.08 | 0.18 | -0.02 | -0.11 | -0.14 | 0.14 | -0.02 | 0.04 | 0.11 |

| | H2 | H3 | H4 | IM1 | IM2 | Exam1 | Exam2 | Exam3 | Cumul |
|---|---|---|---|---|---|---|---|---|---|
| H2 | 1.00 | | | | | | | | |
| H3 | 0.25 | 1.00 | | | | | | | |
| H4 | 0.35 | 0.23 | 1.00 | | | | | | |
| IM1 | 0.16 | 0.14 | 0.07 | 1.00 | | | | | |
| IM2 | 0.23 | 0.13 | 0.27 | 0.24 | 1.00 | | | | |
| Exam1 | 0.08 | 0.10 | 0.06 | 0.23 | -0.01 | 1.00 | | | |
| Exam2 | 0.14 | 0.01 | 0.09 | 0.18 | 0.00 | 0.55 | 1.00 | | |
| Exam3 | 0.10 | 0.11 | 0.12 | 0.16 | -0.03 | 0.46 | 0.43 | 1.00 | |
| Cumul | 0.13 | 0.09 | 0.11 | 0.24 | -0.02 | 0.84 | 0.81 | 0.77 | 1.00 |

Before analyzing the data, the skewness of the responses was examined. After completing factor analysis, the study's observed variables that had originally displayed extreme skewness were not included among the items used in the analysis. Some of the remaining items in the analysis did show a slight skewness in distribution. Several of the variables under *beliefs*, *contextual*, and *grit* showed slight tendencies of left-skew, while *extrinsic motivation* and *anxiety* leaned towards right-skew patterns. Normal distributions were found among variables in *intrinsic motivation, anxiety, habits,* and *grit*. Overall, even though small amount of skewness were observed for some of the variables, future consideration for any impact on the quality of any of these construct's measurement may be required.

**4.2 Measurement Structure**

Validity of a scale refers to the relationship between a theoretical construct and its measure (i.e., whether the behavior of the measure is "consistent with theoretical expectations") (Cronbach, 1955). For this study's exploratory factor analysis as part of scale construction, items with loadings above .40 were retained; items with lower loadings were discarded. If latent constructs did not have any items greater than .40, all items with loadings greater than .32 were kept. The items of the 1st survey that were discarded for the purpose of latent variable construction were still kept as observed variables for the 2nd and 3rd surveys, in order to subsequently test for their stability as non-significant factor loadings. After removing nine items, the factor analysis resulted in 15-items that measured six distinct constructs (with the *motivation* dimension sub-scaled into *extrinsic* and *intrinsic*) for the 1st wave.

Components of this new final scale resulted from some of the following observations: (a) Because item 1 from *habits* addressed procrastination and had a factor loading of .03 when grouped with the other items in the construct, this item did not align in communality with items 2 through 4, since these items concern content studied for the exam. (b) The two strongest loadings (and only ones above. 40) in the *belief* construct came from the incremental theory and entity theory categories of Dweck's Theory of Intelligence Scale (Dweck, 1999). (c) In the *anxiety* construct, the discarded 2nd item did not load well (.16), possibly because it did not address course- or content-related anxiety, where the three retained items of *anxiety* did, and was extremely left-skewed.

**4.3 Model Fit**

To obtain a meaningful test of model fit for the small number of items, several constraints were imposed to over-identify the models. For two-indicator latent constructs this involved: (a) fixing the latent variable's variance to 1, (b) setting both indicators' error terms variances equal to each other, and (3) setting the loading coefficients equal to each other. For the three-indicator latent constructs, the variance of the latent variable was fixed to 1 and two of the loadings were set equal to each other. Table 5 shows the final set of items that were analyzed for this study. The 2nd and 3rd surveys added the *grit* construct, with items 3, 5, and 6 providing sufficient loadings onto one factor. This resulted in 18 items that were able to measure seven distinguishable constructs from both the 2nd survey and 3rd surveys. The factor loadings for many of the retained variables consistently increased through the 2nd and 3rd surveys. Of the seven constructs, *anxiety* and *intrinsic motivation* showed the greatest increase in factor loadings over the three survey waves. *Anxiety*, *intrinsic motivation*, *habits*, and *grit* were measured reasonable well as indicated by factor loadings as high as .5 and .6. The *context* construct did not provide for a reliable, consistent measure, which was most likely a result of these items having been distinctly created by the study's author. Dropping the latent variable, *context*, resulted in six distinct constructs, comprised of non-cognitive indicators, explaining student performance.

TABLE 5
*List of Selected Items and Loading Factors for Non-Cognitive Constructs*

| Factor | Indicator | Standardized Factor Loadings<br>Wave 1 | Wave 2 | Wave 3 |
|---|---|---|---|---|
| Anxiety | 1: Mathematics makes me feel uncomfortable and nervous. | 0.34 (1) | 0.47 (1) | 0.53 (1) |
| | 4: My performance tends to suffer when I have to race against deadlines. | 0.37 (1) | 0.48 (.89) | 0.53 (.45) |
| | 5: The new symbols and Greek letters, in this class, can make me very anxious. | 0.75 (2.4) | 0.64 (1.4) | 0.63 (.77) |
| Beliefs | 1: You have a certain amount of intelligence and you really cannot do much to change it. | 0.48 (1) | 0.31 (1) | 0.51 (1) |
| | 2: You can always greatly change how intelligent you are. | 0.51 (.79) | 0.32 (.15) | 0.55 (.66) |
| Ext Mot | 1: I am doing pre-business (vs. other majors) to have the best opportunities to find a job upon graduation. | 0.38 (1) | 0.46 (1) | 0.43 (1) |
| | 2: I just want to avoid doing poorly, grade-wise, in this class. | 0.35 (.9) | 0.52 (.82) | 0.51 (1.1) |
| | 3: I am very concerned with my parent's thoughts of my grades. | 0.37 (1.4) | 0.37 (.96) | 0.53 (1.6) |
| Int Mot | 1: I would take a Statistics course in future as an elective (non-required) course. | 0.51 (1) | 0.62 (1) | 0.62 (1) |
| | 2: I am more concerned with learning the content of the class than earning a good grade. | 0.50 (1) | 0.68 (.78) | 0.67 (.82) |
| Habits | 2: I have been regularly doing non-graded practice problems for each chapter. | 0.46 (1) | 0.66 (1) | 0.67 (1) |
| | 3: I have sufficiently timed myself when working out problems in preparation for the exam. | 0.42 (.89) | 0.66 (.84) | 0.63 (.61) |
| | 4: I have been consistently reading the required chapters in preparation for the exam. | 0.62 (.84) | 0.63 (.86) | 0.56 (.65) |
| Grit | 3: I am a hard worker. | --- | 0.68 (1) | 0.68 (1) |
| | 5: I am diligent. | --- | 0.72 (1.1) | 0.65 (1.4) |
| | 6: I finish whatever I begin. | --- | 0.47 (.73) | 0.42 (.84) |

*Note. Grit* was not a construct during wave 1. Unstandardized Factor Loading in (parenthesis)

The model fit was evaluated for the measurement and structural models using the chi-square ($\chi^2$), comparative fit index (CFI), and root mean square error of approximation (RMSEA). Less emphasis was placed on the $\chi^2$ statistic, compared with the CFI and RMSEA, as the $\chi^2$ statistic is sensitive to departures from multivariate normality (Yu, 2002). Additionally, it is nearly always large and statistically significant for larger sample sizes (Hsieh, 2012), which this study employed. CFI, RMSEA and TLI partially correct for sample size and model complexity, providing a better assessment of model fit (Hsieh, 2012). TLI values over .90 are considered acceptable (Bentler & Hu, 1999). CFI statistics greater than .90 and RMSEA values less than .06 were considered good model fits (Yu, 2002). An RMSEA below .10 is considered an adequate fit (Yu, 2002).

The model fit of the confirmatory procedure attempted to confirm what this study's exploratory procedure produced. The model fit for each of the latent constructs in each of their respective one-factor model analysis is displayed in Table 6. Model fit statistics for all constructs across three waves are displayed in Table 6.

TABLE 6
*Model Fit of Latent Variables*

| | Wave 1 | | | | | | | Wave 2 | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $X^2$ | P-Val | df | CFI | RMSE | $R^2$ | n | $X^2$ | P-Val | df | CFI | RMSE | $R^2$ | n |
| Anxiety (3) | 0.01 | 0.91 | 1 | 1.00 | 0.00 | 0.61 | 486 | 0.39 | 0.54 | 1 | 1.00 | 0.00 | 0.56 | 498 |
| Beliefs (2) | 1.95 | 0.16 | 1 | 0.97 | 0.04 | 0.40 | 424 | 19.48 | 0.00 | 1 | 0.00 | 0.00 | 0.20 | 488 |
| Ext Motiv (3) | 0.05 | 0.82 | 1 | 1.00 | 0.00 | 0.48 | 466 | 0.40 | 0.53 | 1 | 1.00 | 0.00 | 0.44 | 484 |
| Int Motiv (2) | 0.06 | 0.81 | 1 | 1.00 | 0.00 | 0.39 | 484 | 8.10 | 0.00 | 1 | 0.94 | 0.12 | 0.61 | 538 |
| Habits (3) | 0.30 | 0.58 | 1 | 1.00 | 0.00 | 0.53 | 480 | 2.52 | 0.11 | 1 | 0.99 | 0.05 | 0.69 | 569 |
| Grit (3) | ---- | ---- | ---- | ---- | ---- | ---- | ---- | 0.34 | 0.56 | 1 | 1.00 | 0.00 | 0.70 | 561 |

*Note*. *Grit* was not a construct during wave 1. (2) = 2-indicator model, (3) = 3-indicator model.

TABLE 6 (cont'd)

| | Wave 3 | | | | | | |
|---|---|---|---|---|---|---|---|
| | $X^2$ | P-Val | df | CFI | RMSE | $R^2$ | n |
| Anxiety (3) | 21.87 | 0.00 | 1 | 0.85 | 0.13 | 0.45 | 453 |
| Beliefs (2) | 8.49 | 0.00 | 1 | 0.83 | 0.04 | 0.40 | 481 |
| Ext Motiv (3) | 0.21 | 0.64 | 1 | 1.00 | 0.00 | 0.45 | 369 |
| Int Motiv (2) | 3.96 | 0.05 | 1 | 0.97 | 0.00 | 0.08 | 463 |
| Habits (3) | 10.93 | 0.00 | 1 | 0.94 | 0.15 | 0.66 | 427 |
| Grit (3) | 3.54 | 0.06 | 1 | 0.98 | 0.07 | 0.65 | 423 |

The measurement model helped to determine that each construct had large enough standardized loadings and understand if there was sufficient model fit before evaluating the impact of this study's structural models. The 3-indicator constructs mostly show good model fit with the exception of *anxiety* in wave 3 as its RMSEA (.13) and CFI (.85) did not reach an adequate fit. Additionally, even though *habits* had a good CFI value (.94) in wave 3, its RMSEA (.15) did not achieve an adequate fit. The 2-indicator model of *beliefs* did not have a good model fit in wave 3 with a CFI value of .83. Similarly, *intrinsic motivation*, which had a CFI of .94 in wave 2, did not have adequate fit based on its RMSEA(.12, ).

Following validation of the construction of the study's six latent variables, each scale was checked to determine the extent to which the scale produces scores that are free from measurement error (Hinkin, 1998). The latent constructs' scale reliability, based on their predicted factor scores, can be seen in Table 7. These results confirm the pattern of the factor loadings (as shown in Table 5) over the course of the three waves. Of the three surveys, waves 2 & 3 showed the strongest reliability measurements. Overall, *habits, grit, intrinsic motivation* and *anxiety* provided for the most reliable assessments of their respective constructs. Of these four constructs, only intrinsic motivation is a 2-indicator construct. The other 2-indicator construct, *beliefs*, lost some reliability in the 2$^{nd}$ wave, which confirmed what was reported in the confirmatory factor analysis. It should be noted and taken into consideration, however, that the reliability measure of a 2-indicator construct can be seen as inappropriate and meaningless (Eisinga, 2013). Though none of these constructs' exhibit internal consistency above the psychometric benchmark value of .80 (DeVellis, 1991), or the commonly accepted value of .70 for research, anxiety, study habits, grit and for wave 2 and 3 intrinsic motivation do achieve one previous study's suggestion of .50 as an acceptable cutoff point (Bowling, 2002). In consideration of the social sciences and the small number of items representing each construct, all latent variables (with the exception of *context*) will be used for the structural equation modeling and latent class analysis portions of this study.

TABLE 7
*Reliability Scale (Cronbach's Alpha)*

| Factor | Wave 1 | Wave 2 | Wave 3 |
|---|---|---|---|
| Anxiety | 0.47 | 0.54 | 0.58 |
| Beliefs | 0.39 | 0.31 | 0.46 |
| Extrinsic Mot | 0.30 | 0.39 | 0.43 |
| Intrinsic Mot | 0.40 | 0.61 | 0.59 |
| Study Habits | 0.51 | 0.69 | 0.65 |
| Grit | --- | 0.66 | 0.61 |

*Note. Grit* was not a construct during wave 1.

## 4.4 Structural Equation Modeling

In an attempt to answer this study's 2nd (constructs' impact on exam performance) and 3rd (stability of explanatory power over time) research questions, several SEM models were constructed in order to determine the explanatory strength of the latent variables on exam performance. As these models were tested over the three different waves, determination of stability and predictive consistency was explored. Before constructing each SEM model, a measurement model in which all latent constructs were correlated was fitted to test for model specification errors. The determined significant correlations were then modeled in each of the respective SEM models. The maximum likelihood (ML) method was used for estimating the models. Using a list-wise deletion method, the number of observations varied by model.

Model 1 tested the latent variables (in the first wave "(1)") that had measures of reliability higher than .40, namely *anxiety* (A), *habits* (H) and *intrinsic motivation* (I). Upon checking for correlations between all the constructs in Model 1, there were significant correlations found involving *habits* and *anxiety*, as well as *habits* and *intrinsic motivation*. Incorporating these correlations into the model produced the loading coefficients seen in Figure 4's SEM. The exogenous variables all have a unidirectional (recursive) path to the endogenous variable, *exam1*. When the loading coefficients are standardized, the three latent variables all met expected theory and had a positive effect on *exam1* performance, with *anxiety* having the greatest weight (.48). As the *anxiety* items were reverse coded, a positive coefficient suggests that students with low anxiety (high scale score) performed better on the 1st exam than those with high anxiety. More precisely, if a student's *anxiety* improves by 1 standard deviation (1.4 scale units), it can be expected that their performance on the 1st exam would improve by .48 standard deviations, or 6.71 points (1 standard deviation = 13.98 percentage points), holding all other exogenous variables constant. Considering reverse coding, the negative correlation between *anxiety* and *habits* suggests the less anxiety a student has, the lower their study habits. This may be the result of students who feel adequately prepared from past mathematics coursework may be less anxious and, thus, are not inclined to study as much. Figure 4 shows the direct effects of the exogenous variables' coefficients in their standardized form.

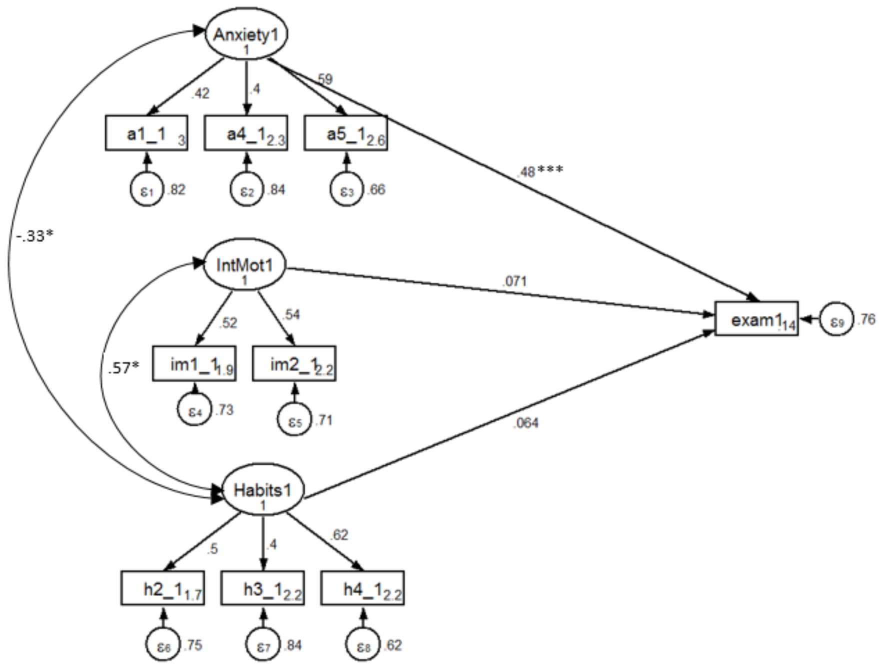


*Figure 4: Wave1, 3-Factor SEM Model (Anxiety, Intrinsic Motivation, Habits. Exam1)*
*Note:* *p-value < .05. ***p-value < .001.

The exogenous variables from Model 1 were able to explain 24.2% of the variation of 1st exam performance (Table 8). Though *anxiety* was found to be statistically significant in this model (p < .001), *habits* (p = .11) and *intrinsic motivation* (p = .12) were not. As seen in Table 8, the CFI and TLI indices (.899 & .855) for Model 1 came close to satisfying the established criteria for an adequate fitting model. The Chi-Square criterion was not satisfied ($\chi^2$ = 47.38, p = .004). According to RMSEA indices, this model had a fair fit (RMSEA = .050).

TABLE 8
*SEM Models (3-Factor & 4-Factor)*

| Model | Endogenous | Exogenous | $X^2$ (df) | P-Val | RMSEA | $R^2$ | CFI | TLI |
|---|---|---|---|---|---|---|---|---|
| 1 | Exam 1 | (1) A***, H, I | 47.38 (25) | 0.004 | 0.050 | 0.242 | 0.899 | 0.855 |
| 2 | | (1) A*, H*, I, B* | 106.11 (47) | 0.000 | 0.060 | 0.322 | 0.765 | 0.725 |

*Note*. Model 1: n=360, Model 2: n=346. A = anxiety, H = habits, I = intrinsic motivation, B = beliefs, (1) = wave 1. *p-val < .05, ***p-val < .001

The *beliefs (B)* variable, with a Cronbach's alpha (α) of .39 (Table 7), was added as a 4th exogenous variable to those used in Model 1. Upon checking for correlations between all the constructs in a measurement model, it was found that correlations between *anxiety* and *habits*, *intrinsic motivation* and *habits*, as well as *beliefs* and *intrinsic motivation* existed. This assisted in Model 2 being able to explain 32.2% of the variation of the first exam, though the CFI and TLI indices both decreased substantially from Model 1. *Beliefs* also became statistically significant in this model (p=.03). This suggests that a student with a growth mindset (vs. fixed mindset) will have a better performance on their 1st exam. The addition, the *beliefs* construct may help to explain why some model misfit partially results from a construct(s) with poor reliability measurement. Overall in the 1st wave, a low amount of *anxiety* was the strongest and only statistically significant exogenous variable appearing in both models in predicting performance on the 1st exam. Lastly, the initially observed correlation between *beliefs* and *intrinsic motivation* was lower and statistically no longer significant (p-value = .071) when tested with this SEM (Model 2), for specification purposes it was still included.

To test for consistency and stability of these constructs' predictive power over time (3rd research question), the same non-cognitive constructs in Model 1 (*anxiety, habits, intrinsic motivation*) were retained as exogenous variables for the 2nd wave in order to determine their effects on the 2nd exam. Again testing for model specification errors, it was found that only a significant correlation existed between *anxiety* and *habits*. Incorporating this correlation into the model again produced positive loading coefficients. While *intrinsic motivation* became significant (p = .03), a*nxiety* lost its significance (p = .07), and *habits* remained non-significant. Similar to the 1st wave, all the loading coefficients had a positive effect on exam performance, with *intrinsic motivation* having the greatest impact on exam performance. This suggests that as a student measures higher in intrinsic motivation, their performance on the 2nd exam would be expected to increase. As seen in Table 9, Model 3's power decreased by over one-third from that of Model 1, as it declined to .155 (from $R^2$ = .242).

TABLE 9
*SEM Models with Varying Factors*

| Model | Endog | Exogenous | $X^2$ (df) | P-Val | RMSE | $R^2$ | CFI | TLI |
|---|---|---|---|---|---|---|---|---|
| 1 | Exam 1 | (1) A***, H, I | 47.380 (25) | 0.004 | 0.050 | 0.242 | 0.899 | 0.855 |
| 2 | | (1) A*, H*, I, B* | 106.110 (47) | 0.000 | 0.060 | 0.322 | 0.765 | 0.725 |
| 3 | Exam 2 | (2) A, H, I* | 45.137 (28) | 0.021 | 0.067 | 0.155 | 0.845 | 0.801 |
| 4 | | (2)A**,H,I**,G* | 67.104 (54) | 0.109 | 0.031 | 0.188 | 0.949 | 0.938 |
| 5 | Exam 3 | (3) A**, H*, I | 41.538 (23) | 0.010 | 0.049 | 0.095 | 0.941 | 0.916 |
| 6 | | (3) A***,H, I, G* | 103.886 (55) | 0.000 | 0.051 | 0.140 | 0.886 | 0.863 |

*Note*. A = anxiety, H = habits, I = int motivation, B = beliefs, E = ext motivation, G = grit, (1) = wave 1, (2) = wave 2, (3) = wave 3. *p-val < .05, **p-val < .01, ***p-val < .001. Model 1: n=360, Model 2: n=346, Model 3: n=472, Model 4: n=256, Model 5: n=332, Model 6: n=343.

As mentioned earlier, the 2nd survey added the dimension of *grit* as a potential indicator of student performance. As seen in Tables 5 and 6, *grit (G)* had relatively strong internal consistency and the 2nd highest reliability score for each respective wave, so it was included. In keeping with the premise of including the constructs that had the strongest measures of reliability (Table 7) in each wave, Model 4 included *grit* along with *anxiety*, *habits* and *intrinsic motivation* in a recursive SEM model (Figure 5). Upon checking for correlations between all the constructs, *intrinsic motivation* and *habits*, as well as *anxiety* and *habits* had significant correlations and were included in the model to avoid specification error. It helped to improve the model to acceptable fit criteria (CFI = .949, TLI = 0.938). Adding this more reliable construct increased fit from Model 3 as well as increased explanation of variation ($R^2$ = .188). The positive correlation between *intrinsic motivation* and *habits* suggest that the greater the interest a student has in learning the subject matter, an improvement in study habits will result. Again, considering reverse coding, the negative correlation between *anxiety* and *habits* suggests the less anxiety a student has, the lower their study habits. Except for *habits*, the other three constructs were statistically significant. A positively loaded coefficient of *grit* would explain that the more resiliency a student has, the better their exam performance. Figure 5 shows the direct effects of the exogenous variables' coefficients.

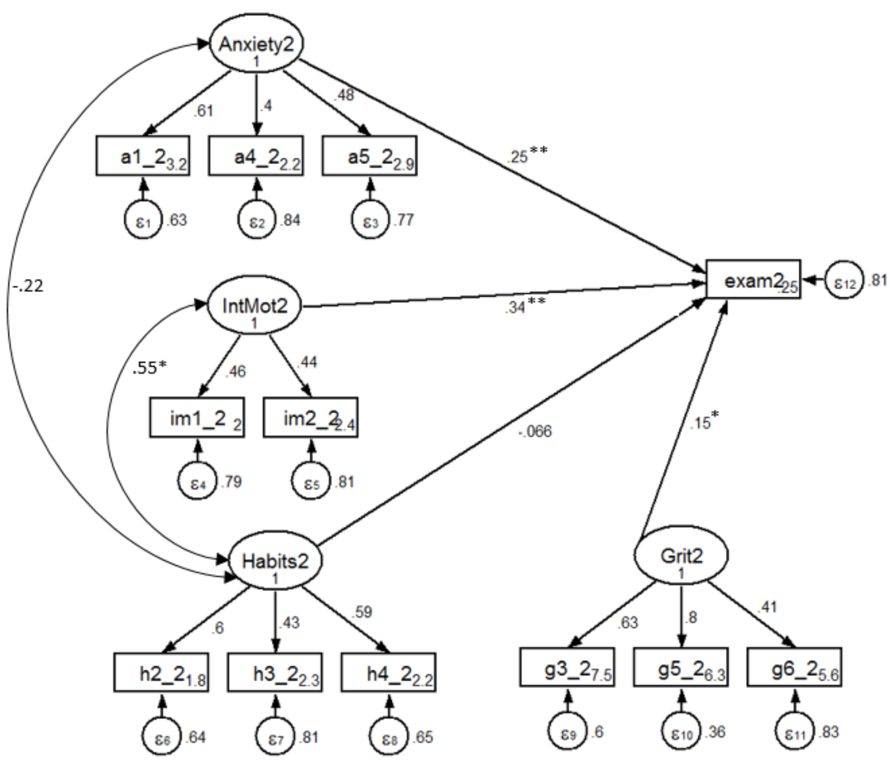


*Figure 5. Wave 2: 4-Factor SEM Model (Anxiety, Intrinsic Motivation, Habits, Grit. Exam2)*
*Note:* *p-value < .05. **p-value < .01.

In wave 3, Models 5 & 6 included latent variables with the strongest measures of reliability in order to predict *exam3* performance. These models combined to be the best fitting models of all three waves (CFI = .941, TLI = 0.916 & CFI = .886, TLI = 0.863). In model 5, the exogenous variables *anxiety*, *habits*, and *intrinsic motivation*, continued its decline of predictive power from waves 1 and 2 ($R^2$ = .095). Correlations between *anxiety* and *habits*, as well as *intrinsic motivation* and *habits* were included as they were found to be significant. As *intrinsic motivation's* highest reliability score (α = .59) came in wave 3, it also remained as one of the stronger constructs, though lost its individual significance (from wave 2). *Anxiety* became significant again ($p < .01$), though the degree of *anxiety's* significance did not approach that of wave 1. *Anxiety's* loading coefficients in waves 2 and 3's (.21 & .27) 3-factor models had dropped almost half from its wave 1 contribution (0.48). *Habits* emerged for the first time as significant. In both models 5 & 6, only *anxiety* had significance ($p < .01$). Positive loading coefficients of the significant variables continued to meet theoretical expectations.

Adding *grit* to this 3-factor model increased $R^2$ (.140) (Model 6) with the resulting loading shown in Figure 6. Correlations between *anxiety* and *habits*, as well as *intrinsic motivation* and *habits* were included as they were found to be significant. Figure 6 shows the direct effects of the exogenous variables' coefficients in their standardized form.

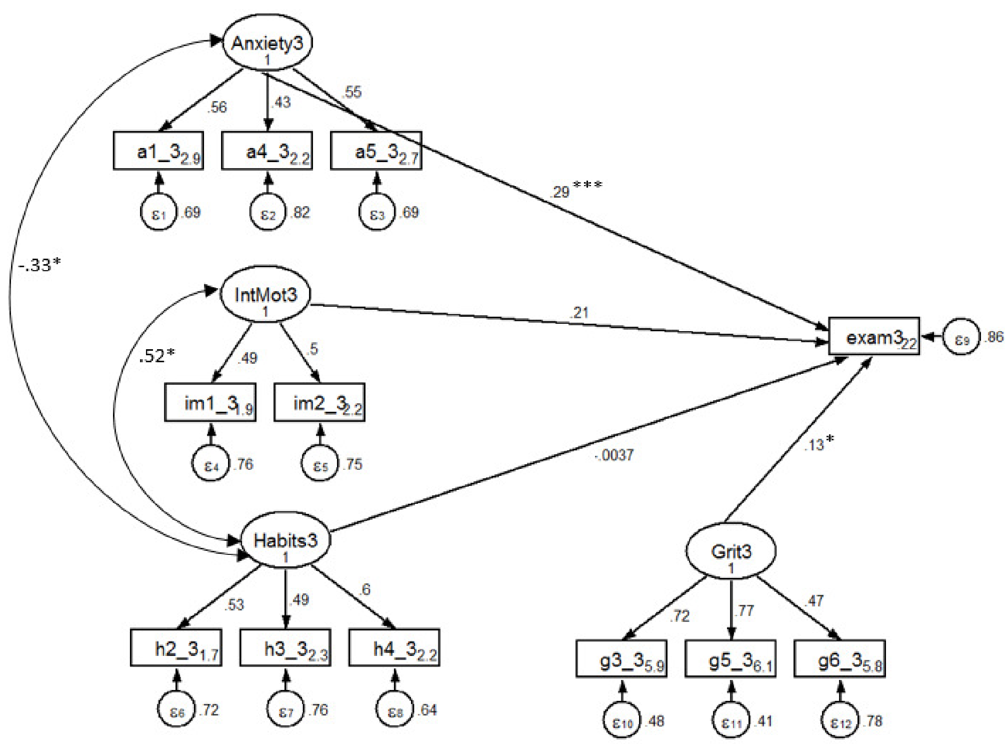


*Figure 6. Wave 3, 4-Factor SEM Model (Anxiety, Intrinsic Motivation, Habits, Grit. Exam 3)*
*Note:* *p-value < .05. ***p-value < .001.

While the overall predictive power of the study's models decreases over time, the relative impact of different non-cognitive indicators changes over the three waves. In the 3-factor constructs, *anxiety* contributed the most to explaining variation of performance for exams 1 and 3, while *intrinsic motivation* was the most significant for exam 2. These results confirm previous theories concerning *anxiety* as the most relevant non-cognitive predictor of mathematics exam performance (Hembree, 1990). *Habit's* consistency on impact of academic performance in this study, although not statistically significant in this study, mirrors previous research's findings of consistent results (Entwistle, 1971). However, when modeled with *grit,* some of *habit's* impact tended to be absorbed (wave 3) by the other exogenous variables. In every wave, the 4-factor model explained the most variation in the endogenous variable. When waves 2 and 3 added the 4$^{th}$ factor *grit*, it substantially increased its predictive power over each wave's 3-factor model which confirms its addition as an important factor on 2$^{nd}$ exam performance.

Continuing to test for the stability and consistency of the latent constructs, this study next explored how wave 1 constructs of the 4-factor model would predict *later* exams and cumulative exam scores. It was hypothesized that a path of exogenous variables from the first wave would result in predictive effects separately on exam2 (*exam2),* exam 3 (*exam3*) and cumulative exam performance (*cumul*). Again accounting for model specification errors, correlations involving *habits* and *anxiety*, *habits* and *intrinsic motivation*, as well as *beliefs* and *intrinsic motivation* wereincluded in these models as they were found to be significant.

When *exam2* and *exam3* was respectively designated as an endogenous variable (Models 8 & 9) (Table 10) with wave 1 exogenous variables, the explanatory power dropped ($R^2 = .091$ and $R^2 = 0.142$) from that of Model 7 ($R^2 = .322$). The CFI and TLI indices also noted a drop in model fit which might be expected when using indicators for exogenous variables from an earlier wave (1) to predict later waves' exam performance. Interestingly, student responses to the survey items from wave 1 ($R^2 = .142$) offered a greater explanation of *exam3's* variation than did the wave 3 items ($R^2 = .095$).

TABLE 10

*4- Factor SEM Models (Wave 1 Constructs)*

| Model | Endog | Exogenous | $X^2$ (df) | P-Val | RMSE | $R^2$ | CFI | TLI |
|---|---|---|---|---|---|---|---|---|
| 7 | Exam 1 | A***, H, I, B* | 106.110 (47) | 0.000 | 0.060 | 0.322 | 0.765 | 0.725 |
| 8 | Exam 2 | A*, H, I, B | 91.290 (47) | 0.000 | 0.060 | 0.091 | 0.779 | 0.703 |
| 9 | Exam 3 | A**, H*, I, B | 108.620 (47) | 0.000 | 0.058 | 0.142 | 0.743 | 0.718 |
| 10 | Cumul | A***, H*, I, B | 89.160 (47) | 0.000 | 0.060 | 0.244 | 0.801 | 0.726 |

*Note*. n= 346. A = anxiety, H = habits, I = int motivation. B = beliefs. *p-val < .05, **p-val < .01, ***p-val < .001.

When the endogenous variable was substituted with *cumul* (Model 10), the explanation of variation ($R^2 = .244$) showed a marked improvement from Models 8 & 9. As Table 10 shows, the varying degree of *anxiety's* significance seems to reflect the predictive power of each model. When *anxiety* is not as significant in a model, its $R^2$ declines accordingly. Loading coefficients with *cumul* as the endogenous variable (Figure 7), exhibited direct effects mostly consistent with expectations. Lastly, although the correlation between *beliefs* and *intrinsic motivation* had varying p-values between .056 and .07 (non-significant) in the above models, for specification purposes they were still included.

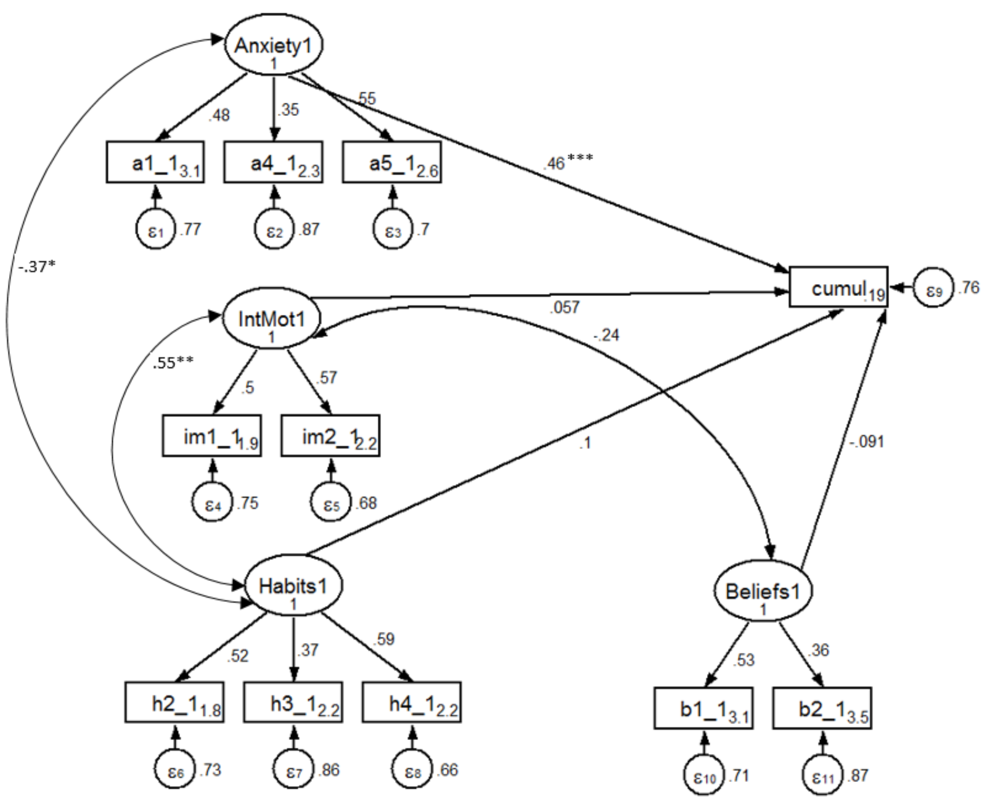


*Figure 7. Wave 1: 4-Factor SEM Model (Anxiety, Intrinsic Motivation, Habits, Beliefs. Cumul)*
*Note:* *p-value < .05. ***p-value < .001.

Since the constructs' explanatory power lost strength over time, this study's latent constructs may become less relevant as the semester progresses. The possibility also exists that a weaker predictive power is a result of 1) omitted variables not measured in this study or 2) varying sample sizes over the waves. The weaker predictive power cannot be seen as a result of the models' fits as they were generally increasing over the three waves. This suggests the changing significance of constructs and/or omitted variables might be more a plausible explanation for loss in predictive power. Figure 7 shows the direct effects of the exogenous variables' coefficients in their standardized form.

As previous literature suggests, cognitive predictors account for between one-third to one-half of explanatory power of academic performance (Lavin, 1965). A cognitive variable was investigated in this study (Table 11) to help understand the unique explanatory power of the non-cognitive constructs. As a single exogenous variable, the cognitive item, *prev,* explained 19.9% (Model 11) of variation of the endogenous variable, *cumul. Prev* was next added to this study's recurring non-cognitive construct of *anxiety*, *intrinsic motivation* and *habits* in predicting separately *exam1* and *cumul.* When added with the non-cognitive constructs, this model was able to explain 30.4% of the *cumul's* variation, with only the *anxiety* construct significant. When adding the cognitive variable, *prev,* to 4-factor non-cognitive set, the model fit decreased substantially (CFI = .630, TLI = 0.593 & CFI = .707, TLI = 0.672).

TABLE 11
*3- Factor SEM Model (Wave 1 Constructs)*

| Model | Endogenous | Exogenous | $X^2$(df) | P-Val | RMSE | $R^2$ | CFI | TLI |
|---|---|---|---|---|---|---|---|---|
| 11 | Cumul | Prev | 0.00 (0) | --- | 0.000 | 0.199 | 1.000 | 1.000 |
| 12 | Exam1 | Prev***,A***, H,I,B | 182.87 (59) | 0.00 | 0.077 | 0.327 | .630 | .593 |
| 13 | Cumul | Prev***,A***,H, I,B | 160.18 (59) | 0.00 | 0.071 | 0.304 | .707 | .672 |

*Note*. Prev = previous math grades, A = anxiety, H = habits, I = intrinsic motivation, B = beliefs. "---" = model not identified. ***p-value < .001. Model 11: n=581, Models 12 & 13: n = 344.

### 4.5 Latent Class Analysis

Having established the relatively strongest latent constructs that best explain exam performance in the pre-business statistics class, this study next investigated ways to accurately identify the students most likely to underperform through Latent Class Analysis (LCA). The first step in the LCA process was to recode the six ordered Likert responses into three ordered values. For the software's algorithm to identify classes, the number of possible combinations of observed variable responses must be reduced for more stability (Magidson & Vermunt, 2002). Originally on a scale from 1 to 6, the observed variables within each construct (for each student) were averaged together and then re-coded on a scale from 1 to 3, with each increment representing a proportionally higher advancement in positive non-cognitive engagement. Any construct with a value greater than 4.35 was coded as a response of "3" (strong) for the student. A construct with a value between 2.68 and 4.34 was coded as "2" (medium). Finally, any construct with a value between 1 and 2.67 was coded as "1" (weak). For this study's LCA, the data from the latent constructs of the 1st wave (vs. 2nd or 3rd) was chosen since an intervention occurring as early in the semester as possible requires data collection from the earliest wave.

As the underlying construct in LCA is a categorical variable, student performance (i.e., 1st exam scores, cumulative exam scores, etc.) had to be categorized from their continuous, standardized form. The standardized values of each performance measure were first sorted in descending order. Based on the sorting, these values were then placed into one of five quintiles. Each standardized value was then coded according to their respective quintile. The value of "5" was given for any observation in the top quintile, with "1" assigned to the lowest quintile group. In this multinomial form, student performance measures could act as covariates in the LCA. Additionally, a new measure of student performance was created for this LCA. As the study's author has witnessed many students successfully recover after a poor first exam performance, the variable *bounce-back (BB)* was created to measure the difference between a student's standardized 1$^{st}$ exam score and their standardized cumulative exam score. This would allow for investigation into whether or not some non-cognitive constructs help explain student turnaround following the 1$^{st}$ exam. The resulting values of this new covariate *BB* were sorted and coded in the same manner as the other two covariates. In order to test for LCA classification stability, each of these three covariates were separately tested with this study's multi-factor LCA models.

Educators aim to identify underperforming students and mitigate the path toward continuously poor results. Hence, a primary consideration in determining ways to group combinations of latent constructs for LCA involves the idea of malleability. This involved determining a specific combination of factors that can illuminate the need for outside-the-classroom development of mindset and resilience. A 3-factor LCA was hypothesized involving *beliefs, intrinsic motivation* and *grit.* Research provides support in the malleability of these constructs by way of successful intervention (Dweck, 2008) (Kamenetz, 2016). Though *grit* was not an established construct in the 1st survey, for purpose of this latent class analysis, an average of this construct's items from the 2nd and 3rd waves were used and then coded in the same manner as the other non-cognitive constructs.

Because the goal of LCA aims to determine the smallest number of latent classes, T, the process began by fitting successive models to the data, each time adding another dimension by increasing the number of classes by one. This continued until the simplest model (ie. smallest number of classes) provided an adequate fit. The 3-factor model of *beliefs*, *grit* and *intrinsic motivation* was first tested with the covariate *exam1*. As the optimal number of classes was not known *a priori*, models between one and four classes were estimated (i.e., a 1-class model, a 2-class model, a 3-class model and a 4-class model). Table 12 shows that with $L^2(H_0) = 116.9$ (df = 29) and p = .66, the amount of association (non-independence) that exists in these data can be explained by chance, so the null model cannot be rejected in favor of T > 1 classes. Testing the same 3-factor model with the covariate *cumul* provided similar results (p = .18) (Table 13).

TABLE 12
*3 Factor LCA Model (with Covariate Exam1)*

| Factors | Covariate | Classes | LL | BIC(LL) | Npar | L² | df | P-Val |
|---|---|---|---|---|---|---|---|---|
| B, G, I | Exam 1 | 1-Class | -1280.9 | 2598.0 | 6 | 116.9 | 124 | 0.66 |
| B, G, I | Exam 1 | 2-Class | -1277.1 | 2621.0 | 11 | 109.3 | 119 | 0.73 |
| B, G, I | Exam 1 | 3-Class | -1272.2 | 2642.0 | 16 | 99.6 | 114 | 0.83 |
| B, G, I | Exam 1 | 4-Class | -1269.3 | 2667.0 | 21 | 93.7 | 109 | 0.85 |

*Note*. B = beliefs, G = grit, I = intrinsic motivation. n = 447.

TABLE 13
*3 Factor LCA Model (with Covariate Cumul)*

| Factors | Covariate | Classes | LL | BIC(LL) | Npar | L² | df | P-Val |
|---|---|---|---|---|---|---|---|---|
| B, G, I | Cumul | 1-Class | -1280.9 | 2598.0 | 6 | 138.3 | 124 | 0.18 |
| B, G, I | Cumul | 2-Class | -1274.2 | 2616.0 | 11 | 124.9 | 119 | 0.34 |
| B, G, I | Cumul | 3-Class | -1271.5 | 2641.0 | 16 | 119.4 | 114 | 0.34 |
| B, G, I | Cumul | 4-Class | -1267.5 | 2663.0 | 21 | 111.6 | 109 | 0.41 |

*Note*. B = beliefs, G = grit, I = intrinsic motivation. n = 447.

Testing the 3-factor model *beliefs*, *grit* and *intrinsic motivation* with the covariate *BB* (difference in *exam1* and *cumul* score) produced different results than did the previous two covariates. As seen in Table 14, since $L^2(H_0) = 151.822$ (df = 124) and p < .05, the amount of association existing in this data is too large to be explained by chance, so the null model must be rejected in favor of T > 1 classes. The 2-class model was next considered and resulted in a 12.2% reduction in its baseline model ($L^2$ from 151.8 to 133.2), providing an adequate overall fit ($p > .05$). This model's ability to produce distinct classes suggest that the malleable factors of a person's *beliefs*, *grit* and *intrinsic motivation* may play a strong role in their ability to bounce back from a poor first exam performance.

TABLE 14
*3-Factor LCA Model (with Covariate Diff)*

| Factors | Covariate | Classes | LL | BIC(LL) | Npar | $L^2$ | df | P-Val | Class. Err. |
|---|---|---|---|---|---|---|---|---|---|
| B, G, I | Diff | 1-Class | -1280.9 | 2598.4 | 6 | 151.8 | 124 | 0.04 | 0.00 |
| B, G, I | Diff | 2-Class | -1271.6 | 2610.3 | 11 | 133.2 | 119 | 0.18 | 0.19 |
| B, G, I | Diff | 3-Class | -1264.3 | 2626.2 | 16 | 118.6 | 114 | 0.37 | 0.28 |
| B, G, I | Diff | 4-Class | -1261.9 | 2651.9 | 21 | 113.7 | 109 | 0.36 | 0.34 |

*Note*. B = beliefs, G = grit, I = intrinsic motivation. n = 447.

When classification of cases is based on modal assignment (the class having the highest membership probability) the proportion of cases that are expected to be misclassified is reported as classification error (Magidson & Vermunt, 2005). The 2-class model incurred a classification error of .19. The class probabilities of the two-class model are shown in Table 15 and are arranged from largest to smallest. For Class 1, the probability was .55, indicating that 55% of students belonged here. For Class 2, the category probability was .45, meaning 45% (the remaining students) were assigned to this group.

TABLE 15

*Profile Output, 2-Class Model*

| | Class 1 | Class 2 |
|---|---|---|
| Indicators | 0.5541 | 0.4459 |
| Beliefs | | |
| 1 | 0.0268 | 0.0724 |
| 2 | 0.3751 | 0.5122 |
| 3 | 0.5981 | 0.4158 |
| Mean | 2.5713 | 2.3437 |
| Grit | | |
| 1 | 0.2391 | 0.4003 |
| 2 | 0.3505 | 0.3522 |
| 3 | 0.4105 | 0.2475 |
| Mean | 2.1714 | 1.8473 |
| Int Mot | | |
| 1 | 0.2143 | 0.3216 |
| 2 | 0.6244 | 0.5843 |
| 3 | 0.1622 | 0.0947 |
| Mean | 1.9481 | 1.7737 |
| Covariate diff | | |
| 1 | 0.0345 | 0.3937 |
| 2 | 0.0906 | 0.2989 |
| 3 | 0.1986 | 0.1897 |
| 4 | 0.3247 | 0.0881 |
| 5 | 0.3517 | 0.0295 |
| Mean | 3.8686 | 2.0608 |

Analogous to factor analysis where names are assigned to factors based upon an examination of the factor loadings, names may be assigned to the latent classes based upon estimated conditional probabilities. Like factor loadings, the conditional probabilities provide the measurement structure that defines the latent classes (Magidson & Vermunt, 2004). As Class 1 all had construct means (and covariate mean) higher than Class 2, Class 1 was thus labeled the "high positive non-cognitive engagement" group, compared to *low* for Class 2. Those in the *high* category had a mean approaching the 4$^{th}$ quintile (3.87) of exam improvement (since 1$^{st}$ exam), while *low* had a mean approximating the 2nd quintile. The amount of this model's classification error (.19) suggests that 19% of the students assigned to the *high* class may be misclassified and supposed to belong in the *low* class. This suggests a false positive where the amount of students in the *low* class may be under-identified. As it concern's classification error, under-identification of at-risk students that would benefit from an intervention is a valid concern and will be discussed later.

The parameters in Table 15 are in terms of column percentages. For example, 59.8% of the *high* students had a response of "3" (strong) in the *beliefs* construct as opposed to 2.7%, in the same class, who had a response of "1" (weak). Another interpretation can be made in comparison of the classes. As it concerns the covariate *BB*, a *high* student is almost 12 times (.352/.029) more likely to be in the top quintile than a student from the *low* classification. The association for such a greater likelihood lies in each respective class' responses to constructs of *beliefs*, *grit* and *intrinsic motivation.* For example, a student from the *high* class is 1.7 (.411/.248) times more likely to have a response of "3" on the grit construct than a student from the *low* class. Viewing these probabilities graphically can be seen in Figure 8. On the horizontal axis are the various constructs and covariate. The vertical axis contains scaled values (between 0 to 1) of a variable's mean. This scaled value is accomplished by subtracting the lowest observed value from the cluster-specific means and dividing the results by the range (Magidson &Vermunt, 2005).

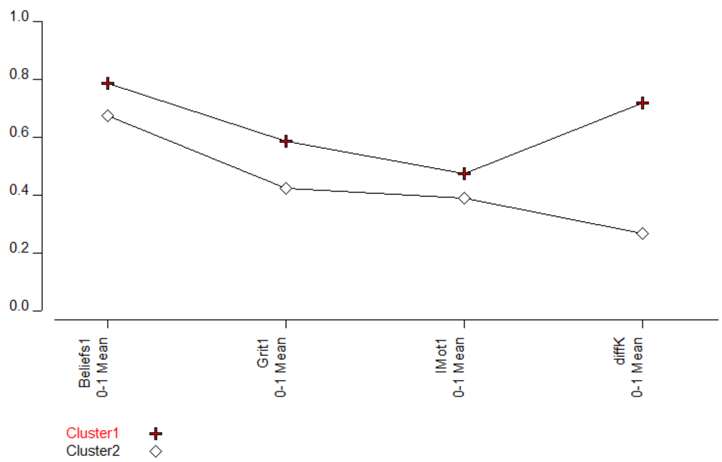


*Figure 8 3-Factor, 2-Class LCA Model (Belief, Grit, Int Motiv. BB)*

A 2nd 3-factor model also involved an element of malleability as its grouping was based on (a) an instructor's ability to accommodate shortcomings indicated within the construct and (b) the strength of construct measurement (as determined earlier in this study). The three chosen factors of *anxiety, habits*, and *intrinsic motivation* all met the criterion for measurement strength and potential instructor accommodation. A few means by which instructors can address poor performance include developing learning strategies (*habits*), expressing enthusiasm for the subject matter to encourage student engagement, and providing a strong foundation for understanding Greek symbols (*anxiety*).

Following the same classification approach as the previous 3-factor model, this 3-factor model's pairing with each of the three covariates had optimal results with a 3-class model (Table 16). When *exam1*, *cumul* or *BB* acted as the covariate, all the models' parameters were significant ($p<.05$). This suggests the LCA classification of this 3-factor model represents consistency and stability across any survey wave. Table 17 shows results of the three factors with the covariate, *cumul*. All p-values here (<.05) indicate that each construct contributes in a significant way towards the ability to discriminate between the classes. The $R^2$ values indicate how much each construct's variance is explained by this 3-class model (eg., 36.8% of the variance of the anxiety construct is explained).

TABLE 16
*3-Factor LCA Models*

| Factors | Covariate | Classes | All Param's Signif? | LL | BIC(LL) | Npar | $L^2$ | df | P-Val | Class. Err. |
|---|---|---|---|---|---|---|---|---|---|---|
| I, A, H | Exam1 | 3 | Yes | -1479.4 | 3057.6 | 16 | 121 | 114 | 0.31 | 0.17 |
| I, A, H | Cumul | 3 | Yes | -1482.6 | 3064.1 | 16 | 125 | 114 | 0.23 | 0.16 |
| I, A, H | Diff | 3 | Yes | -1489.7 | 3078.3 | 16 | 123 | 114 | 0.27 | 0.14 |

*Note*. I = intrinsic motivation, A = anxiety, H = habits. n = 483

TABLE 17
*Parameter Summary 3-Factors, 3-Class LCA*

| Models For Indicators | Class 1 | Class 2 | Class 3 | Wald | P-Val | $R^2$ |
|---|---|---|---|---|---|---|
| Intrinsic motivation | 0.574 | 1.559 | 2.133 | 15.60 | 0.000 | 0.214 |
| Anxiety | 1.401 | 0.833 | 0.568 | 7.370 | 0.025 | 0.368 |
| Habits | 0.537 | 0.260 | 0.277 | 9.940 | 0.007 | 0.085 |
| | | | | | | |
| Model For Clusters | Class 1 | Class 2 | Class 3 | Wald | P-Val | |
| Intercept | 0.102 | 0.997 | 1.099 | 4.932 | 0.085 | |
| | | | | | | |
| Covariates | Class 1 | Class 2 | Class 3 | Wald | P-Val | |
| Cumulative | 0.307 | 0.371 | 0.065 | 10.373 | 0.006 | |

Still considering the 3-factor model with the covariate, *cumul*, Table 18 reveals 66.8% of students were assigned to Class 1 and 23.9% of students were assigned to Class 2. Class 1 all had construct means (and covariate mean) higher than Class 2, and was thus labeled "*high* positive non-cognitive engagement" group versus "*low*" for Class 2. Again, the *high* class had the highest membership probability where a classification error rate of .16 (Table 16) in this class suggests a false positive which in turn may under-identify the number of at-risk students.

With mixed results, only 9.2% of students comprised Class 3. A closer inspection of this small class reveals that the students performed relatively well (*cumul* mean: 3.0), even though they had low scores on *anxiety* and *habits*. It is possible the high score on *intrinsic motivation* buoys this select group's exam performance. Because Class 3 does not likely follow the ordered ranking of the other two classes, this group may be most aptly labeled "high intrinsic motivation".

TABLE 18
*Profile Summary 3-Factor, 3-Class Model*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| Class Size | 0.668 | 0.239 | 0.092 |
| Indicators | | | |
| Intrinsic motivation | | | |
| 1 | 0.225 | 0.458 | 0.007 |
| 2 | 0.674 | 0.513 | 0.304 |
| 3 | 0.102 | 0.029 | 0.689 |
| Mean | 1.877 | 1.571 | 2.682 |
| Anxiety | | | |
| 1 | 0.107 | 0.695 | 0.627 |
| 2 | 0.383 | 0.267 | 0.315 |
| 3 | 0.510 | 0.038 | 0.059 |
| Mean | 2.403 | 1.343 | 1.432 |
| Habits | | | |
| 1 | 0.213 | 0.466 | 0.471 |
| 2 | 0.343 | 0.337 | 0.336 |
| 3 | 0.444 | 0.197 | 0.193 |
| Mean | 2.231 | 1.731 | 1.721 |
| Covariates | | | |
| Cumulative | | | |
| 1 | 0.099 | 0.362 | 0.162 |
| 2 | 0.162 | 0.271 | 0.232 |
| 3 | 0.223 | 0.195 | 0.206 |
| 4 | 0.227 | 0.103 | 0.243 |
| 5 | 0.289 | 0.072 | 0.163 |
| Mean: | 3.445 | 2.253 | 3.006 |

Whereas Table 18 expresses the parameters in terms of column percentages, Table 19 re-expresses them in terms of row percentages. Interpretation of some notable probability means' results (Table 19) for this 3-class model follow:

1) A student designated to the *high* class is 11.3 times (.858/.076) more likely to be in the top quintile of cumulative exam performance than a student in the *low* class.

2) A student designated to the *high* class is 34.2 times (.957/.028) more likely to have a response of "3" (very positive) in the *anxiety* construct than a student in the *low* class.

3) A student designated to the *low* class is 6.8 times (.517/.076) more likely to be in the lowest quintile of cumulative exam performance than in the top quintile.

4) A student designated to the *low* class is 8.6 times (.422/.049) more likely to have a response of “1” (strong), in the *intrinsic motivation* construct, than a response of “3” (weak)

TABLE 19
*Probability Means for 3-Class Model*

| | Class 1 | Class 2 | Class 3 |
|---|---|---|---|
| Prob Means | 0.670 | 0.239 | 0.092 |
| Indicators | | | |
| Intrinsic motivation | | | |
| 1 | 0.580 | 0.422 | 0.003 |
| 2 | 0.755 | 0.205 | 0.046 |
| 3 | 0.494 | 0.049 | 0.462 |
| Anxiety | | | |
| 1 | 0.242 | 0.567 | 0.196 |
| 2 | 0.740 | 0.177 | 0.084 |
| 3 | 0.969 | 0.028 | 0.014 |
| Habits | | | |
| 1 | 0.471 | 0.389 | 0.142 |
| 2 | 0.692 | 0.212 | 0.099 |
| 3 | 0.810 | 0.142 | 0.045 |
| Covariates | | | |
| Cumulative | | | |
| 1 | 0.393 | 0.517 | 0.089 |
| 2 | 0.561 | 0.332 | 0.110 |
| 3 | 0.790 | 0.217 | 0.088 |
| 4 | 0.774 | 0.124 | 0.112 |
| 5 | 0.863 | 0.076 | 0.066 |

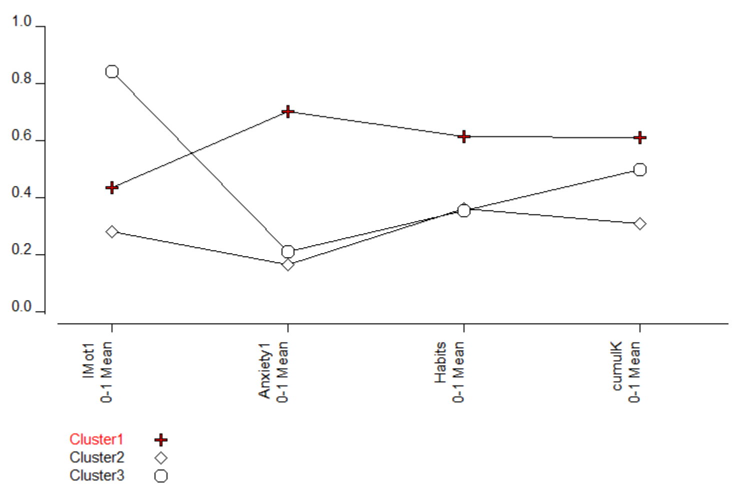


*Figure 9 3-Factor, 3-Class LCA Model (Int Motiv, Anxiety, Habits. Cumul)*

As identification of classes with more indicators, i.e. more information, can possibly benefit the quality of the LCA, the last model considered for LCA was a 4-factor model. As this study's sample size was large enough (Wurpts & Geiser, 2014) to handle an additional factor, the relatively reliable construct *grit* joined the other malleable factors *anxiety*, *habits* and *intrinsic motivation.* Following the same classification approach as the previously-constructed LCA models, this 4-factor model's pairing with each of the three covariates had optimal results with a 2-class model (Table 20). When *exam1*, *cumul* or *BB* acted as the covariate, all the models' parameters were significant ($p < .05$). This suggests the LCA classification of this 4-factor model represents stability across the three different performance measures. Table 21 shows results of the four factors with the covariate, *cumul*. The $R^2$ values indicate that *anxiety* (.02) and *intrinsic motivation* (.06) have the least explained variance.

TABLE 20
*Parameter Summary of 4-Factor LCA Models (varying Covariate)*

| Factors | Covariate | Class | All Param's Signif? | LL | BIC | Npar | $L^2$ | df | P-Val | Class. Error |
|---|---|---|---|---|---|---|---|---|---|---|
| A, I, H, G | Exam1 | 2 | Yes | -1930.2 | 3945.2 | 14 | 416.6 | 386 | 0.14 | 0.17 |
| A, I, H, G | Cumul | 2 | Yes | -1927.8 | 3940.5 | 14 | 381.9 | 386 | 0.55 | 0.16 |
| A, I, H, G | Diff | 2 | Yes | -1928.5 | 3934.6 | 14 | 420.5 | 386 | 0.11 | 0.14 |

*Note*. I = intrinsic motivation, A = anxiety, H = habits, G = grit. n = 462.

TABLE 21
*4-Factor LCA Model (Covariate: Cumul)*

| <u>Models for Indicators</u> | Class 1 | Class 2 | Wald | P-Val | $R^2$ |
|---|---|---|---|---|---|
| Anxiety | 0.200 | 0.200 | 3.486 | 0.042 | 0.020 |
| Int Motiv | 0.473 | 0.473 | 5.013 | 0.025 | 0.061 |
| Habits | 0.659 | 0.659 | 11.503 | 0.001 | 0.171 |
| Grit | 0.863 | 0.863 | 8.153 | 0.004 | 0.228 |
| | | | | | |
| <u>Model for Clusters</u> | Class 1 | Class 2 | Wald | P-Val | |
| Intercept | 1.076 | 1.076 | 7.562 | 0.006 | |
| | | | | | |
| <u>Covariates</u> | Class 1 | Class 2 | Wald | P-Val | |
| Cumul | 0.173 | 0.173 | 6.209 | 0.013 | |

Still considering the 4-factor model with the covariate, *cumul*, Table 22 shows 73.4% of students were assigned to Class 1 and 26.7% of students were assigned to Class 2. Class 1 all had construct means (and covariate mean) lower than Class 2, and was thus labeled "low positive non-cognitive engagement" group versus *high* for Class 2. With this model's classification error of .16 (Table 20), it can be interpreted that 16% of the students assigned to the *low* class may belong in the *high* class. This suggests that the amount of students in the *low* class may be over-identified. Over-identification of at-risk students is preferable to under-identification as there would be no disservice for a student to receive assistance that did not need it. Interpretation of the more notable results from this 4-factor/2-class model follows:

TABLE 22
*4-Factor, 2-Class LCA Model*

| | Class 1 | Class 2 |
|---|---|---|
| Class Size | 0.7337 | 0.2663 |
| Indicators | | |
| Anxiety | | |
| 1 | 0.360 | 0.235 |
| 2 | 0.340 | 0.331 |
| 3 | 0.300 | 0.435 |
| Mean | 1.939 | 2.200 |
| Int Motiv | | |
| 1 | 0.311 | 0.126 |
| 2 | 0.593 | 0.618 |
| 3 | 0.096 | 0.257 |
| Mean | 1.785 | 2.131 |
| Habits | | |
| 1 | 0.443 | 0.095 |
| 2 | 0.347 | 0.278 |
| 3 | 0.209 | 0.627 |
| Mean | 1.766 | 2.532 |
| Grit | | |
| 1 | 0.406 | 0.045 |
| 2 | 0.392 | 0.245 |
| 3 | 0.203 | 0.710 |
| Mean | 1.797 | 2.665 |
| Covariates | | |
| Cumul | | |
| 1 | 0.184 | 0.087 |
| 2 | 0.219 | 0.136 |
| 3 | 0.227 | 0.187 |
| 4 | 0.180 | 0.262 |
| 5 | 0.191 | 0.329 |
| Mean | 2.974 | 3.609 |

1) A student designated to the *low* class is 4.67 times (.443/.095) more likely to have a response of "1" in the *habits* category than a student in the *high* class.

2) A student designated to the *high* class is 3.5 times (.710/.203) more likely to have a response of "3" in the *grit* construct than a student in the *low* class.

3) A student designated to the *high* class is 1.6 times {(.262+.329)/(.180+.191} more likely to be above the 3rd quintile of cumulative exam performance than a student in the *low* class.

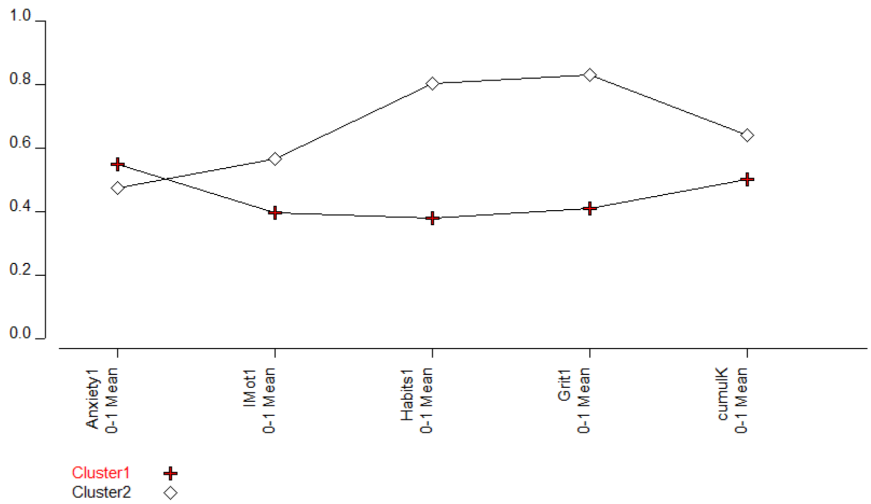


*Figure 10* ***4-Factor, 2-Class LCA Model (Anxiety, Int Motiv, Habits, Grit. Cumul)***

The results of this study's LCA provide evidence for the relevance that these constructs have in distinguishing students who may be prone to underperforming in the business statistics course.

## 5. Discussion

Early studies in education have shown that cognitive indicators explain just under half of the variation in academic achievement (Lavin, 1965). This leaves ample opportunity to explore non-cognitive factors influencing academic performance. Particularly for the 2nd business statistics course, identifying and emphasizing non-cognitive factors that positively contribute to exam performance can spur students toward future success in the business program. This study explored such factors and aimed to provide knowledge for building a diagnostic tool distinguishing students who may or may not perform well in the course. To do so, this study investigated 1) whether the items in the survey measured non-cognitive factors in student performance well, and relatedly, whether the study could include new items and validate a new scale with internal consistency and reliability complementing items established in previous literature; 2) whether latent constructs explain exam performances, and how the strength of their reliability influences model fit; 3) whether the results are stable and consistent over time; and 4) whether the study allows for the identification of students likely to underperform, given this study's selected malleable factors.

This study uncovered several important findings. The first finding bridged the study's measurement and SEM analysis. Although the measurements for the SEM's latent constructs were not optimal, the constructs' impact on predicting exam performance became apparent. When accounting for factors found most reliable in the 1st part of the study (*anxiety, habits and intrinsic motivation*), their impact on the 1st exam and cumulative scores was indubitable. These factors (during wave 1) explained approximately almost one-quarter of two endogenous variable's variation (*exam1*, *cumul*). These results are consistent with previous literature, reaffirming the validity of the study's choice of constructs. Additionally, when correlations were considered between constructs, *intrinsic motivation* and *habits* as well as *intrinsic motivation* and *beliefs* were significant. These results also confirm previous literature's finding of strong associations between the respective factors.

A second finding involved the study's strongest constructs' decline (approximately 50%) in explanatory power over time, possibly resulting from the strength of the factors' individual significances changing over the course of the semester. SEM analysis found that *anxiety* lost its predictive power over the course of the semester, although it remained significant. The descriptive bivariate relationships (See Appendix E) of *anxiety's* observed variables during waves 2 and 3 also confirmed this result, as each indicator's relationship to their respective exam period fell by over half. Similar to the reasoning that prior mathematics courses' performance (*prev*) falls in relational strength over the course of a semester (as it did in this study), perhaps uncertainty prior to taking their 1st exam of the course, and prior knowledge that the rigorous material involves mathematics/statistics, played a key role in student performance. However, as students better understood the class' statistics exams, anxiety's effect on performance decreased.

As *anxiety* fell by over half in its individual significance, the overall explanatory power of the SEM multi-factor models involving *anxiety* also fell by more than half. As *anxiety* decreased in its significance, *intrinsic motivation* emerged as an important contributor. Perhaps as the rigors of a semester continue, earnest understanding of the course material gains importance in exam performance. Because *intrinsic motivation* does not highly prioritize short-term considerations, its significance manifests when long-term performance results are considered. This factor's ascendency, however, did not offset *anxiety's* decline in the overall explanatory power of the models over the three waves. The changing significance of these constructs, over the course of the semester, suggests a strong argument for their malleability, particularly as it concerns the possibility that instructors can improve student performance. Stemming the decline in the predictive power of the SEM models, over the 2nd and 3rd waves, involved the addition of *grit* to the latent-construct set. When *grit* was added to the SEM models in these two waves, it resulted in a 4-factor model which improved upon the respective wave's 3-factor model. As the predictive power of the model decreased over the three waves, the measurement strength of the constructs' reliability (Cronbach's alpha) continued to increase. As the factors became more reliable, an improved model fit mostly resulted. This, however, did not affecy the models' decreasing coefficient of determination over the three waves. Relevant factors early in the semester became less impactful as the semester progressed. This may also suggest some relevant factors for explaining exam performance may be missing during the later waves of the semester.

The last finding concerns the strength of the study's constructs to diagnose underperforming students. Based on a review of the previous literature, this study appears to be the first in using latent class analysis to investigate the non-cognitive factors associated with undergraduate performance in a business statistics class. The LCA still successfully classified student groups using three and four factor models of the 1st wave's strongest constructs (*anxiety*, *intrinsic motivation*, *habits* and *grit*), although this wave had its lowest reliability measures of all three waves. This success occurred when paired with any exam-performance covariate. Perhaps with improved measurement constructs, future research to test and develop LCA models will result in even more precise classifications of student subgroups. Reasoning would suggest that improved, more reliable, constructs would result in a lower classification error. There is no harm in over-identifying at-risk students, though there may be harm in under-identification of students needing assistance. Successful classification should, in turn, help diagnose at-risk candidates in this challenging undergraduate statistics class and suggest interventions that address mindsets, learning strategies, resiliency, and anxiety to significantly improve student performance in the classroom.

### 5.1 Limitations

This study's results must be interpreted in light of several limitations. As with many self-response surveys, some biases may have influenced the results. Although students were always encouraged to answer questions honestly and to the best of their ability, since there were no wrong answers, certain responses may not have conformed to this request. As the surveys replaced the usual means of recording lab attendance, a student could not be penalized for their response. Also, although questions were not ordered randomly, the questions were not presented in the same order for every survey. The collection of survey responses for 25% of the students fell on a Monday, while the remaining responses were collected on a Wednesday, a few days closer to the exam. Perhaps collecting data so close to the exam days inflated the relations between variables. Another bias may have occurred with the observations that were ultimately employed in the analysis using STATA's ML (listwise deletion) method. A response bias may be present when only considering participants who provided responses to all items. The students who completed all responses might be significantly different from those whose responses were not analyzed. As it concerns number of surveys completed, N's varied across the analyses which may have ultimately had an affect on the results. Additionally, some model misfit may have been a product of occasional reactive survey development, which kept other indicators in this analysis from surfacing. Given the exploratory nature of this study, no mediation or moderation models were tested and hence the study cannot provide any information on potentially important indirect effects on student exam performance. Lastly, though the study's results were somewhat consistent with previous theory, the generalizability of these results is limited to this particular business statistics class.

**5.2 Future Work**

At the business school, students normally evaluate the teaching and the instructor. Evaluations of students generally occurs with evaluation of performances on exams, i.e., cognitive factors. It is quite possible that some analysis of these evaluations have involved latent class analysis. However, conducting evaluations on students involving non-cognitive factors may be particularly helpful in diagnosing at-risk students. With increased precision of future LCA models as a diagnostic tool, interventions can be tailored to target the student subgroups that would benefit most. Brief interventions (e.g., 2 to 10 hours) can significantly impact students' mindsets and learning strategies, resulting in better academic performance (DOE, 2013). Empirically based mindset interventions include activities that explicitly teach students to have a "growth mindset" and also affirm students' personal values to clarify why they are investing their efforts in their studies (intrinsic motivation) (DOE, 2013). Intervention to reduce anxiety in statistics (and improve performance) can be accomplished by encouraging students to discuss their fears (Dillon, 1982), gathering data from the students themselves, tasking students to perform simple calculations (Schacht & Stewart, 1991) and being sensitive to students' concerns (Pan & Tang, 2005). "Emphasizing control of negative emotional responses to math stimuli (rather than merely additional math training) will be most effective in revealing a population of mathematically competent individuals, who might otherwise go undiscovered" (Lyons and Sian L. Beilock, 2011). Behavioral treatments and cognitive-behavioral methods in treating mathematics-related anxiety showed a collective mean improvement of .57 on exam performance (Hembree, 1990). Where study habits are concerned, group learning activities can result in a deep-learning approach (Hall, 2004), which may be investigated during the lab session that students attend in the campus' Collaborative Learning Space.

## 6. Conclusion

This study aimed to (a) propose and test models using non-cognitive factors that explain exam performance among pre-business students in a business statistics class, and (b) diagnose those students susceptible to struggling in the course. Although there is substantial research on this study's chosen variables, it is important for the University's business school to understand how these variables affect student performance on exams for a business statistics course. This study identified the more influential non-cognitive factors on exam performance, with a subsequent latent class analysis affirming those results. Efforts to learn from this study and further development of future instruments for analysis may offer better exploration of factors determining academic success. In addition, this study's findings can be used to complement existing literature and serve as the foundation for future research. Hopefully, this study provides a meaningful contribution to students in challenging courses, as well as to instructors and educators in developing a diagnostic tool allowing successful intervention of high-risk candidates. Early identification and intervention of these candidates will likely result in a stronger performance, both in this business statistics class as well as upper-level coursework, and benefit both student and educator.

Appendix A – Supplementary Tables

*Anova: Single Factor (Exam1)*

| SUMMARY | | | | |
|---|---|---|---|---|
| *Groups* | *Count* | *Sum* | *Average* | *Variance* |
| Column 1 | 396 | 26334 | 66.5 | 146.7975 |
| Column 2 | 132 | 8952 | 67.81818 | 174.6232 |
| Column 3 | 166 | 11032 | 66.45783 | 167.4012 |

ANOVA

| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
|---|---|---|---|---|---|---|
| Between Groups | 189.4729 | 2 | 94.73647 | 0.603446 | 0.547212 | 3.008757 |
| Within Groups | 108481.8 | 691 | 156.9925 | | | |
| | | | | | | |
| Total | 108671.3 | 693 | | | | |

*Anova: Single Factor (Cumul)*

| SUMMARY | | | | |
|---|---|---|---|---|
| *Groups* | *Count* | *Sum* | *Average* | *Variance* |
| Column 1 | 396 | 79002 | 199.5 | 146.7975 |
| Column 2 | 132 | 26640 | 201.8182 | 174.6232 |
| Column 3 | 166 | 33342.4 | 200.8578 | 167.4012 |

ANOVA

| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
|---|---|---|---|---|---|---|
| Between Groups | 608.5226 | 2 | 304.2613 | 1.938062 | 0.144765 | 3.008757 |
| Within Groups | 108481.8 | 691 | 156.9925 | | | |
| | | | | | | |
| Total | 109090.4 | 693 | | | | |

## Appendix B

1st Survey of Social Psychological Attributes and Behavior

**Q1:** My overall average grade in: Last H.S. math class, MAT109, MAT116 & BNAD276.

**A1**. Mathematics makes me feel uncomfortable and nervous.
**A2.** I see myself as someone who is emotionally stable and not easily upset.
**A3.** I feel in control of my time & that I can use time the way I want to in this class.
**A4.** My performance tends to suffer when I have to race against deadlines.
**A5.** The new symbols and greek letters, in this class, can make me very anxious.

**B1.** You have a certain amount of intelligence and you really cannot do much to change it.
**B2.** You can always greatly change how intelligent you are.
**B3.** The resources in this class are not sufficient to obtain the grade I desire.
**B4.** What I learned in last H.S. Math class, MAT109, MAT116 & BNAD276 will greatly determine how successful I am in this course.
**B5.** Very high levels of math are consistently found in this class.

**C1.** I feel like I have a strong social support network in U of A / Tucson to help me through any difficult times.
**C2.** I plan on attending the preceptor review session.
**C3.** In recent weeks, I have been working at a job ______________ per week.
**C4**. I have been accepted into the business school.

***IM1.** I would take a Statistics course in future as an elective (non-required) course.
***IM2.** I am more concerned with learning the content of the class than earning a good grade.

***EM1**. I am doing pre-business (vs. other majors) to have the best opportunities to find a job upon graduation.
***EM2.** I just want to avoid doing poorly, gradewise, in this class.
***EM3.** I am very concerned with my parent's thoughts of my grades.

**H1**. On average, I turn in my weekly case: __________________ .
**H2.** I have been regularly doing non-graded practice problems for each chapter.
**H3.** I have sufficiently timed myself when working out problems in preparation for the exam.
**H4.** I have been consistently reading the required chapters in preparation for the exam.

*Originally in this 1st survey, Intrinsic & Extrinsic Motivation were under the construct of "Motivation.

## Appendix C

2nd Survey of Social Psychological Attributes and Behavior

**A1**. Mathematics makes me feel uncomfortable and nervous.
**A2.** I see myself as someone who is emotionally stable and not easily upset.
**A3.** I feel in control of my time & that I can use time the way I want to in this class.
**A4.** My performance tends to suffer when I have to race against deadlines.
**A5.** The new symbols and greek letters, in this class, can make me very anxious.
**A6.** I often feel overwhelmed by the level of math required for this course.

**B1.** You have a certain amount of intelligence and you really cannot do much to change it.
**B2.** You can always greatly change how intelligent you are.
**B3.** The resources in this class are not sufficient to obtain the grade I desire.
**B5.** Very high levels of math are consistently found in this class.
**B6.** I don't believe only some people are good at math – I believe everyone is capable of learning math
**B7.** Even if I do poorly on an exam, I believe I can improve my score dramatically on next exam.

**C1.** I feel like I have a strong social support network in U of A / Tucson to help me through any difficult times.
**C2.** I plan on attending the preceptor review session.
**C3.** In recent weeks, I have been working at a job _______________ per week.
**C4**. I have been accepted into the business school.

**IM1.** I would take a Statistics course in future as an elective (non-required) course.
**IM2.** I am more concerned with learning the content of the class than earning a good grade.
**IM3.** I am doing pre-business (vs. other majors) to have the knowledge and skills to perform well in a job upon graduation.
**IM4.** In a class like this, I prefer course material that arouses my curiousity, even if it is difficult to learn.
**IM5.** It is important for me to learn what is being taught in this class.

**EM1**. I am doing pre-business (vs. other majors) to have the best opportunities to find a job upon graduation.
**EM2.** I just want to avoid doing poorly, gradewise, in this class.
**EM3.** I am very concerned with my parent's thoughts of my grades.
**EM4.** I am only taking this course because it is required.

**H1**. On average, I turn in my weekly case: ___________________ .
**H2.** I have been regularly doing non-graded practice problems for each chapter.
**H3.** I have sufficiently timed myself when working out problems in preparation for the exam.
**H4.** I have been consistently reading the required chapters in preparation for the exam.
**H5.** I usually plan my week's work in advance, either on paper on in my head.
**H6.** I need to be in the right mood before I can study effectively
**H7**. Even when I do poorly on an exam, I try to learn from my mistakes

**G1.** Setbacks don't discourage me.
**G2.** I have achieved a goal that took years of work
**G3.** I am a hard worker
**G4.** I have difficulty maintaining my focus on projects that take more than a few months to complete.
**G5.** I am diligent.
**G6.** I finish whatever I begin.

Appendix D

3rd Survey of Social Psychological Attributes and Behavior

**A1**. Mathematics makes me feel uncomfortable and nervous. (slight left skew)
**A2.** I see myself as someone who is emotionally stable and not easily upset.
**A3.** I feel in control of my time & that I can use time the way I want to in this class.
**A4.** My performance tends to suffer when I have to race against deadlines. (slight right skew)
**A5.** The new symbols and greek letters, in this class, can make me very anxious. (normal)
**A6.** I often feel overwhelmed by the level of math required for this course.

**B1.** You have a certain amount of intelligence and you really cannot do much to change it. (left skew)
**B2.** You can always greatly change how intelligent you are. (slight left skew)
**B3.** The resources in this class are not sufficient to obtain the grade I desire.
**B5.** Very high levels of math are consistently found in this class.
**B6.** I don't believe only some people are good at math – I believe everyone is capable of learning math
**B7.** Even if I do poorly on an exam, I believe I can improve my score dramatically on next exam.

**C1.** I feel like I have a strong social support network in U of A / Tucson to help me through any difficult times. (left skew)
**C2.** I plan on attending the preceptor review session.
**C3.** In recent weeks, I have been working at a job ______________ per week. (left skew)
**C4**. I have been accepted into the business school.

**IM1.** I would take a Statistics course in future as an elective (non-required) course. (slight right skew)
**IM2.** I am more concerned with learning the content of the class than earning a good grade(normal)
**IM3.** I am doing pre-business (vs. other majors) to have the knowledge and skills to perform well in a job upon graduation.
**IM4.** In a class like this, I prefer course material that arouses my curiousity, even if it is difficult to learn.
**IM5.** It is important for me to learn what is being taught in this class.

**EM1**. I am doing pre-business (vs. other majors) to have the best opportunities to find a job upon graduation. (right skew)
**EM2.** I just want to avoid doing poorly, gradewise, in this class. (right skew)
**EM3.** I am very concerned with my parent's thoughts of my grades. (slight right skew)
**EM4.** I am only taking this course because it is required.
**EM5.** My only objective in this class is to earn the highest possible grade I can.

**H1**. On average, I turn in my weekly case: __________________ .
**H2.** I have been regularly doing non-graded practice problems for each chapter. (slight right skew)
**H3.** I have sufficiently timed myself when working out problems in preparation for the exam. (normal)
**H4.** I have been consistently reading the required chapters in preparation for the exam. (normal)
**H5.** I usually plan my week's work in advance, either on paper on in my head.
**H6.** I need to be in the right mood before I can study effectively
**H7**. Even when I do poorly on an exam, I try to learn from my mistakes

**G1.** Setbacks don't discourage me. (normal)
**G2.** I have achieved a goal that took years of work
**G3.** I am a hard worker (slight left skew)
**G4.** I have difficulty maintaining my focus on projects that take more than a few months to complete.
**G5.** I am diligent. )Slight left skew)
**G6.** I finish whatever I begin. (Slight left skew)

## Appendix E – Supplementary Tables

*Correlation Matrix: Anxiety*
*(1st Wave)*

| A1 | A1 | A4 | A5 | Exam1 |
|---|---|---|---|---|
| A1 | 1.00 | | | |
| A4 | 0.00 | 1.00 | | |
| A5 | 0.21 | 0.29 | 1.00 | |
| Exam1 | 0.37 | 0.18 | 0.25 | 1.00 |

*Note*. A = anxiety

*Correlation Matrix: Anxiety*
*(2nd Wave)*

| A1 | A1 | A4 | A5 | Eexam2 |
|---|---|---|---|---|
| A1 | 1.00 | | | |
| A4 | 0.23 | 1.00 | | |
| A5 | 0.32 | 0.29 | 1.00 | |
| Exam2 | 0.22 | 0.10 | 0.14 | 1.00 |

*Note*. A = anxiety

*Correlation Matrix: Anxiety*
*(3rd Wave)*

| A1 | A1 | A4 | A5 | Exam3 |
|---|---|---|---|---|
| A1 | 1.00 | | | |
| A4 | 0.29 | 1.00 | | |
| A5 | 0.45 | 0.22 | 1.00 | |
| Exam3 | 0.18 | 0.09 | 0.08 | 1.00 |

*Note*. A = anxiety

References

Ames, C., & Archer, J. (1988). Achievement goals in the classroom: Students' learning strategies and motivation processes. *Journal of Educational Psychology, 80*(3), 260–267.

Ashcraft, M. H., & Ridley, K. S. (2005). Math-anxiety and its cognitive consequences: A tutorial review. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 315–327). New York, NY: Psychology Press.

Aud, S., Hussar, W., Kena, G., Bianco, K., Frohlich, L., Kemp, J., & Tahan, K. (2011). *The condition of education 2011* (NCES 2011-033). Washington, DC: Government Printing Office.

Babcock, P., & Marks, M. (2010). *Leisure College, USA*. University of California. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.412.2624&rep=rep1&type=pdf May, 2010.

Baquiran, L. A. (2011). Habits and attitudes of freshmen students: Implications for academic intervention programs. *Journal of Language Teaching and Research, 2*(5), 1116–1121.

Betz, N. (1978). Prevalence, distribution, and correlates of math anxiety in college students. *Journal of Counseling Psychology, 25*(5), 441–448.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*(1), 246–263.

Blalock, H. M. (1987). Some general goals in teaching statistics. *Teaching Sociology, 15*(2), 164–172.

Bowling, A. (2002). *Research methods in health: Investigating health and health services*. Maidenhead, United Kingdom: Open University Press.

Bratton Yeager, M. (2009). *A cross-validation study of the College Learning Effectiveness Inventory (CLEI): An abstract of a dissertation*. Manhattan, KS: Kansas State University.

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

Canabal, M. E. (1998). College student degree of participation in the labor force: Determinants and relationship to school performance. *College Student Journal, 32*(4), 597–606.

Cao, L. (2012). Differences in procrastination and motivation between undergraduate and graduate students. *Journal of the Scholarship of Teaching and Learning, 12*(2), 39–64.

Cerna, M. A. (2015). Influence of habits on academic performance of international college students in Shanghai. *Higher Education Studies, 5*(4), 42-55.

Chang, W. (2014). *Grit and academic performance: Is being grittier better?* (Doctoral dissertation). Retrieved from University of Miami Scholarly Repository.

Chao, R. C. (2012). Managing perceived stress among college students: The roles of social support and dysfunctional coping. *Journal of College Counseling, 15*(1), 5–21.

Choi, J. N., & Moran, S. V. (2009). Why not procrastinate? Development and validation of a new active procrastination scale. *The Journal of Social Psychology, 149*(2), 195–211.

Church, M. A., Elliot, A. J., & Gable, S. L. (2001). Perceptions of classroom environment, achievement goals, and achievement outcomes. *Journal of Educational Psychology, 93*(1), 43–54.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis*. Hillsdale, NJ: L. Erlbaum Associates.

Conard, M. A. (2006). Aptitude is not enough: How personality and behavior predict academic performance. *Journal of Research in Personality, 40(3)*, 339–346.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78(1)*, 98–104.

Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation, 10*(7), 1-10.

Credé, M. A. (2008). The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science, 3*(6), 425–453.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52(4)*, 281–302.

Cruise, J. R., Cash, R. W., & Bolton, L. D. (1985). Development and validation of an instrument to measure statistical anxiety. *1985 Proceedings of the Section on Statistical Education* (pp. 92–98). Washington, DC: American Statistical Association.

Day, V., Mensink, D., & O'Sullivan, M. (2000). Patterns of academic procrastination. *Journal of College Reading and Learning, 30(2)*, 120–134.

Dayton, C. M., & Macready, G. B. (1988). Concomitant-variable latent-class models. *Journal of the American Statistical Association, 83(401)*, 173–178.

Department of Education. (2013). *Promoting grit, tenacity and perseverance: Critical factors for success in the 21st Century*. Damascus, Maryland: Penny Hill Press.

DeVellis, R. F. (1991). *Scale development*. Newbury Park, NJ: Sage Publications.

Dillon, K. M. (1982). Statisticophobia. *Teaching of Psychology, 9*(2), 117.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*(6), 1087–1101.

Dupeyrat, C., & Mariné, C. (2005). Implicit theories of intelligence, goal orientation, cognitive engagement, and achievement: A test of Dweck's model with returning to school adults. *Contemporary Educational Psychology, 30*(1), 43–59.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press.

Dweck, C. S. (2008). *Mindsets and math/science achievement*. New York, NY: Carnegie Corporation of New York Institute for Advanced Stud. Retrieved from http://opportunityequation.org/teaching-and-leadership/mindsets-math-science-achievement

Dweck, C. S., & Leggett, E. (1988). A social and cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256–273.

Eggens, L., van der Werf, M. P. C., & Bosker, R. J. (2008). The influence of personal networks and social support on study attainment of students in university education. *Higher Education, 55(5)*, 553–573.

Eisinga, R., Grotenhuis, M., & Pelze, B. The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? International Journal of Public Health, 58(4), 637–642.

Entwistle, N. J., & Brennan, T. (1971). The academic performance of students. *British Journal of Educational Psychology, 41(3)*, 258–267.

Fosnacht, K. (2013, March). *Undergraduate coping with financial stress: A latent class analysis.* Bloomington, IN: Center For Postsecondary Research.

Green, S. B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of coefficient alpha as an index of test unidimensionality. *Educational and Psychological Measurement, 37(4)*, 827–838.

Hall, M. R. (2004). Changing the learning environment to promote deep learning approaches in first year accounting students. *Accounting Education, 13*(4), 489–505.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education, 21*(1), 33–46.

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods, 2(1)*, 104–121.

Hsieh, P., Sullivan J., Sass D. A., & Guerra N. (2012). Undergraduate engineering students' beliefs, coping strategies, and academic performance: An evaluation of theoretical models. *The Journal of Experimental Education, 80*(2), 196–218.

Hu, L. T., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76–99). Thousand Oaks, CA: Sage.

John, O. P., & Srivastava, S. (1999). The big-five trait taxonomy: History, measurement, and theoretical perspectives. In *Handbook of personality: Theory and research* (2nd ed, 60-71.). New York, NY: Guilford.

Kamenetz, A. (2016). MacArthur "genius" Angela Duckworth responds to a new critique of grit. *NPR*. Retrieved from http://www.npr.org/sections/ed/2016/05/25/479172868/angela-duckworth-responds-to-a-new-critique-of-grit

Kim, E., Newton, F. B., Downey, R. G., & Benton, S. L. (2010). Personal factors impacting college student success: Constructing College Learning Effectiveness Inventory (CLEI). *College Student Journal, 44(1)*, 1–112.

Laursen, B., & Hoff, E. (2006). Person-centered and variable-centered approaches to longitudinal data. *Merrill-Palmer Quarterly, 52*(3), 601–632.

Lavin, D. E. (1965). *The prediction of academic performance*. New York, NY: Russell Sage Foundation.

Le, H., Casillas, A., Robbins, S. B., & Langley, R. (2005). Motivational and skills, social, and self-management predictors of college outcomes: Constructing the student readiness inventory. *Educational and Psychological Measurement, 65(3)*, 482–508.

Lebold, W. K., & Ward, S. K. (1988, June). *Retention: National and institutional perspectives.* Presented at the American Society for Education Annual Conference, Portland, OR.

Lubke, G., & Muthén, B. O. (2007). Performance of factor mixture models as a function of model size, covariate effects, and class-specific parameters. *Structural Equation Modeling, 14(1)*, 26–47.

Lyons, I. M., & Beilock, S. L. (2011). *Mathematics anxiety: separating the math from the anxiety*. Department of Psychology, University of Chicago, Chicago, IL.

Magidson, J., & Vermunt, J. K. (2002). Latent class cluster analysis. In J. A. Hagenaars & A. L. McCutcheon (Eds.), *Applied latent class analysis* (pp. 89–106). Cambridge: Cambridge University Press.

Magidson, J., & Vermunt, J. K. (2004). Latent class models. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences* (pp. 175–198). Thousand Oaks: Sage Publications.

Magidson, J., & Vermunt, J. K. (2005). *Technical guide to Latent GOLD 4.0: Basic and advance*d. Belmont, MA: Statistical Innovations.

Malik, S. (2015). Undergraduates' Statistics Anxiety: A Phenomenological Study. *The Qualitative Report,* 20(2), 120–133.

Middleton, J., & Toluk, Z. (1999). First steps in the development of an adaptive theory of motivation. *Educational Psychologist, 34*(2), 99–112.

Mji, A. (2009). Differences in university students' attitudes and anxiety about statistics. *Psychological Reports, 104*(3), 737–744.

Nagaoka, J., Farrington, C. A., Roderick, M., Allensworth, E., Seneca, T., Johnson, D. W., & Beechum, N. O. (2013). Readiness for college: The role of noncognitive factors and context. *Voices in Urban Education, 38*, 45–52.

Nonis, S. P. (2006). Where does the time go? A diary approach to business and marketing students' time use. *Journal of Marketing Education, 28*(2), 121–134.

Okun, M. A., Fairholme, C., Karoly, P., Ruehlman, L. S., & Newton, C. (2006). Academic goals, goal process cognition, and exam performance among college students. *Learning and Individual Differences, 16*(3), 255–265.

Ozer, B. U. (2011). A cross sectional study on procrastination: Who procrastinate more? *International Conference on Education, Research and Innovation, IPEDR, 18*, 34–37.

Pan, W., & Tang, M. (2005). Students' perceptions on factors of statistics anxiety and instructional strategies. *Journal of Instructional Psychology, 32*(3), 205–214.

Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology, 92*(3), 544–555.

Reed, W. (1996). S*tudy skills: The key to student success.* Dubuque, IA: Kendall/Hunt Publishing Company.

Richardson, F. C., & Suinn, R. M. (1972). The mathematics anxiety rating scale: Psychometric data. *Journal of Counseling Psychology, 19*(6), 551–554.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130(2)*, 261–288.

Robinson, C. A. (2010). Coaching students to academic success and engagement on campus. *About Campus*, *15(4),* 26–29.

Rummel, R. J. (1970). *Applied factor analysis*. Evanston, IL: Northwestern University Press.

Schacht, S., & Stewart, B. J. (1991). What's funny about statistics? Interactive/user-friendly gimmicks for teaching statistics. *Teaching Sociology, 20*(4), 329–332.

Shipman, S., & Shipman, V. C. (1985). Cognitive styles: Some conceptual, methodological, and applied issues. In E. W. Gordon (Ed.), *Review of research in education* (vol. 12, 229-221). Washington, DC: American Educational Research Association.

StataCorp LP. (2016). *New in Stata 14*. Retrieved from http://www.stata.com/new-in-stata/

Statistical Innovations, Inc. (2016). *Latent GOLD 5.1*. Retrieved from http://www.statisticalinnovations.com/latent-gold-5-1/

Stern, D., & Nakata, Y. (1991). Paid employment among U.S. college students. *Journal of Higher Education, 62*(1), 25–43.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Boston, MA: Allyn and Bacon.

Tinto, V. (2004). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). Chicago, IL: University of Chicago Press.

Tryon, G. (1980). The measurement and treatment of test anxiety. *Review of Educational Research, 50(4)*, 353–372.

VandeWalle, D., Cron, W. L., & Slocum, J. W. (2001). The role of goal orientation following performance feedback. *Journal of Applied Psychology, 86*(4), 629–640.

Van de Water, G. (1996). The effect of part-time work on academic performance and progress: An examination of the Washington State work-study program. In R. Kincaid (Ed.), *Student employment: Linking college and the workplace* (pp. 57–67). Columbia, SC: South Carolina University, National Resource Center for the Freshman Year Experience and Students in Transition.

Velicer, W. F., & Fava, J. L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods, 3*(2), 231–251.

Williams, A. S. (2010). Statistics anxiety and instructor immediacy. *Journal of Statistics Education, 18*(2), 1–18.

Wurpts, I. C., & Geiser, C. (2014). Is adding more indicators to a latent class analysis beneficial or detrimental? Results of a Monte-Carlo study. *Frontiers in Psychology*, *5*, 920-934.

Yu, C. Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes* (Doctoral dissertation). University of California, Los Angeles.

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α, Revelle's β, and Mcdonald's ωH: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika, 70*(1), 1–11.

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2006). Estimating generalizability to a universe of indicators that all have an attribute in common: A comparison of estimators for alpha. *Applied Psychological Measurement, 30(2)*, 121–144.