9-23-2014

Utilizing Traditional Cognitive Measures of Academic Preparation to Predict First-Year Science, Technology, Engineering, and Mathematics (STEM) Majors' Success in Math and Science Courses

Charles K. Andrews
Florida International University, andrewsc@fiu.edu

DOI: 10.25148/etd.FI14110703

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

UTILIZING TRADITIONAL COGNITIVE MEASURES OF ACADEMIC PREPARATION TO PREDICT FIRST-YEAR SCIENCE, TECHNOLOGY, ENGINEERING, AND MATHEMATICS (STEM) MAJORS' SUCCESS IN MATH AND SCIENCE COURSES

A dissertation submitted in partial fulfillment of the

requirements for the degree of

DOCTOR OF EDUCATION

in

HIGHER EDUCATION

by

Charles Andrews

2014

To: Dean Delia C. Garcia
College of Education

This dissertation, written by Charles Andrews, and entitled Utilizing Traditional Cognitive Measures of Academic Preparation to Predict First-Year Science, Technology, Engineering, and Mathematics (STEM) Majors' Success in Math and Science Courses, having been approved in respect to style and intellectual content, is referred to you for your judgment.

We have read this dissertation and recommend that it be approved.

Joy Blanchard

Eric Dwyer

Isadore Newman

Benjamin Baez, Major Professor

Date of Defense: September 23, 2014

The dissertation of Charles Andrews is approved.

Dean Delia C. Garcia
College of Education

Dean Lakshmi N. Reddi
University Graduate School

Florida International University, 2014

ACKNOWLEDGMENTS

I would like to thank my committee members for their support, guidance, and contributions to my study. Drs. Joy Blanchard and Eric Dwyer have provided time and support for the development of my literature review and have offered valuable suggestions for framing my research. Dr. Isadore Newman has been critical to assisting me with the development of my research design and has provided invaluable support for my quantitative analyses. Most of all, I would like to thank my major professor, Dr. Benjamin Baez. He has been my biggest champion throughout this process and has provided me with both scholarly and personal advice that has been instrumental to this process. I cannot thank him enough for all he has done to inspire and motivate me to complete my degree requirements.

In addition to my committee and advisor, I would also like to thank Dr. Douglas Robertson for his ongoing support and encouragement, Marco Gomez from the FIU Academic Advising Technology department for his assistance with setting up my data collection process, and Drs. Dawn Broschard and Connie Boronat from the FIU Office of Retention and Graduation Success for their assistance with my data analysis. I am indebted to Dr. Broschard, in particular, due to her vital role in my process of both understanding and articulating the results of my study.

ABSTRACT OF THE DISSERTATION

UTILIZING TRADITIONAL COGNITIVE MEASURES OF ACADEMIC PREPARATION TO PREDICT FIRST-YEAR SCIENCE, TECHNOLOGY, ENGINEERING, AND MATHEMATICS (STEM) MAJORS' SUCCESS IN MATH AND SCIENCE COURSES

by

Charles Andrews

Florida International University, 2014

Miami, Florida

Professor Benjamin Baez, Major Professor

For the past several years, U.S. colleges and universities have faced increased pressure to improve retention and graduation rates. At the same time, educational institutions have placed a greater emphasis on the importance of enrolling more students in STEM (science, technology, engineering and mathematics) programs and producing more STEM graduates. The resulting problem faced by educators involves finding new ways to support the success of STEM majors, regardless of their pre-college academic preparation. The purpose of my research study involved utilizing first-year STEM majors' math SAT scores, unweighted high school GPA, math placement test scores, and the highest level of math taken in high school to develop models for predicting those who were likely to pass their first math and science courses. In doing so, the study aimed to provide a strategy to address the challenge of improving the passing rates of those first-year students attempting STEM-related courses. The study sample included 1018 first-year STEM majors who had entered the same large, public, urban, Hispanic-serving, research university in the Southeastern U.S. between 2010 and 2012. The research design involved the use of hierarchical logistic regression to determine the significance of utilizing the four independent variables to develop models for predicting success in math and science. The resulting data indicated that the overall model of predictors (which included all four predictor variables) was

statistically significant for predicting those students who passed their first math course and for predicting those students who passed their first science course. Individually, all four predictor variables were found to be statistically significant for predicting those who had passed math, with the unweighted high school GPA and the highest math taken in high school accounting for the largest amount of unique variance. Those two variables also improved the regression model's percentage of correctly predicting that dependent variable. The only variable that was found to be statistically significant for predicting those who had passed science was the students' unweighted high school GPA. Overall, the results of my study have been offered as my contribution to the literature on predicting first-year student success, especially within the STEM disciplines.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I
INTRODUCTION TO THE STUDY

Approximately half of all students who enroll in U.S. colleges and universities have failed to earn a college degree within six years (Stratton, O'Toole, & Wetzel, 2008; Freeman, Hall, & Bresciani, 2007). As a result, retention has become one of the most emphasized aspects of the U.S. higher education system. Colleges and universities have responded by strategizing the best ways to improve student retention and graduation rates. The motivations for this vary but center mostly on the financial and social ramifications of losing students before they are able to complete their college degree. According to Aragon (2000), efforts aimed at increasing retention and learning have created a "challenge for educators to become more competent in the knowledge, skills, abilities, and attitudes that can lead to greater retention" (p. 9).

In an effort to prevent attrition, colleges and universities have begun to implement a number of retention strategies. As both Tinto (1993) and, more recently, Siegel (2011) have noted in their research, college student retention should not merely be a goal but rather a by-product of the educational experiences that institutions provide for their students. According to the ACT's findings on college retention, the most frequent strategies for creating these educational experiences include special programs for first-year students, academic advising, and learning support initiatives (What Works in Student Retention? Fourth National Survey, 2010). In addition to those frequent practices, several colleges and universities have also had success with developing learning communities, implementing special programs for first-generation college students, designing early alert processes, utilizing supplemental instruction, and placing students in appropriate courses based on the results of placement/aptitude tests (Kim, Newton, Downey, & Benton, 2010). The practice of ensuring that students are placed in appropriate courses was particularly important in guiding my research because it prompted questions about whether students who enter higher education institutions are academically prepared for the courses and curricula in which they enroll. According to Freeman et al. (2007), more than half of the variance in institutional retention rates is directly related to attributes of the students rather than to institutional factors. In fact, some of the leading authorities on why college students drop out prior to earning a college degree agree that many students are academically unprepared for the rigors of the academic environments they encounter in college (Daley, 2010; Stratton et al., 2008; Tinto, 1993).

In addition to the emphasis on students' academic preparation, the retention research also highlights the critical nature of the first year in combating college student attrition. Throughout the history of higher education, there has been an increasing focus on the transitions that college students face during their first year. In a recent commentary published in *The Chronicle of Higher Education*, Whelan (2011) pointed out that college freshmen are increasingly overwhelmed by the prospect of starting college. With that in mind, educators have a responsibility to support new students' transition and assist them with the process of establishing a foundation on which to build. According to Stovall (2000), there is a growing body of research to support the claim made by Tinto and other retention experts that the first year is the most critical in determining whether students will persist. One example of this research can be found in Siegel's (2011) *About Campus* article, which highlighted the notion that the first year of college is the most critical year in solving the retention puzzle. The following quote summarizes his thoughts on the ways that the first year can establish a foundation for the future:

> The attitudes, perceptions, and habits students develop in the first year will likely have an enormous influence on their entire college experience. It is critical that institutions take the first year seriously and channel significant resources to curricular and cocurricular structures and academic support services that directly impact first-year students (Siegel, 2011, p. 11).

The first year of college has also received a great deal of attention because of the financial implications for institutions resulting from students leaving college after one year. According to

Schneider (2010), between 2003 and 2008 U.S. states appropriated over $6 billion to colleges and universities to help support the education of students who did not return for their second year. Over that same five-year period, state and federal governments also allocated close to $3 million in grants to assist students who dropped out of college. Kim et al. (2010) highlighted this "bottom line" concern in their research on the factors impacting student success. As they noted, student attrition not only suggests that an institution is lacking in meeting student needs but also impacts its finances through lost tuition dollars. Given the current state of the economy and the increased budget cuts that all colleges and universities have faced in recent years (especially public institutions), it stands to reason that institutions would be concerned with maintaining the revenue generated by tuition and fees. Based on those academic and financial factors related to the transition to college, the emphasis on the importance of the first year became another key aspect of my research.

My review of the literature on retention initiatives that emphasized both individual academic preparation and the importance of the first year led naturally into a consideration of how that research can inform efforts to increase the number of college graduates from the science, technology, engineering, and mathematics (STEM) disciplines. According to Sadler, Sonnert, Hazari, and Tai (2012), preparing high school and college students for careers in STEM is at the forefront of the United States' educational concerns. Thompson and Bolin (2011) shared this sentiment by highlighting how the U.S. has fallen behind other nations and now has one of the lowest rates of graduating students from the STEM disciplines. As part of their research, they also noted that the highest number of STEM dropouts occurs during the first year, reinforcing educators' responsibility to ensure that students have the preparation they need to be successful. My literature review also revealed the importance of pre-college preparation, high school performance, and traditionally relied upon cognitive measures of ability (such as GPA and SAT) on predicting the success of students pursuing STEM degrees (Nicholls, Wolfe, Besterfield-Sacre, Shuman, and Larpkiattaworn, 2007; Sadler et al., 2012; Veenstra, Dey, and Herrin, 2008). For example, Veenstra et al. (2008) found that 38% of the variation in the first-year grade point averages of the engineering majors in their study was attributed to those students' pre-college academic characteristics. In my study, I examined four factors that might help predict if STEM majors at a large, public, urban, Hispanic-serving, research university in the Southeastern U.S. had the academic preparation to succeed in the mathematics and science courses they took during their first year.

Statement of Purpose

The purpose of my quantitative research study was to examine the information that we already know about incoming college freshmen to determine if it helped predict whether STEM majors were likely to pass the math and science courses that they were required to complete during their first year. Given the increased emphasis on students pursuing majors with rigorous math and science curricula, educators may be able to utilize this type of research to diminish the number of students who experience academic difficulty by learning more about predicting their likelihood for success. By determining that certain STEM students are unlikely to pass the first-year math and science courses, new courses can be developed in an effort to provide those students with the additional skills they need to succeed in those courses (or they can be placed into more introductory-level courses that already exist). Some students who possess academic characteristics that make them less likely to succeed might also be encouraged to at least consider pursuing majors outside of STEM. More long term, this information can also be shared with prospective college students to help inform the choices they make while they are still in high school. As Engle and Tinto (2008) pointed out, taking a more rigorous schedule that includes advanced mathematics and science courses will greatly increase the chance of success once students get to college, especially for first-generation college students and those from low income households.

Along those lines, colleges and universities have a responsibility not only to provide greater access to traditionally underrepresented groups but also to work to remove the barriers that have decreased the percentage of low-income and minority students who complete a 4-year degree (Otero, Rivas, & Rivera, 2007; Tinto, 2008). As Williford and Wadley (2008) noted, our goal as educators should be to enable our students to be successful. According to the research, one strategy that may help institutions achieve that is to focus on ensuring that students experience academic success early on. My study aimed to provide data that can assist with predicting that early success. As Johnson (2006) noted, there have been several empirical studies on how a student's "grade performance at the end of the first term has been shown to be the most important factor in college persistence and eventual degree attainment" (p. 927). It is thus important that colleges and universities do more to set students up for success (especially those interested in pursuing STEM majors, with notoriously challenging math and science requirements) by ensuring that they enroll in courses that align with their academic preparation and by providing them with additional resources and support. Therefore, the purposes of my study were to contribute both to the literature on college student success and to provide higher education institutions with an example of how to identify criteria that might predict whether their STEM majors are likely to succeed.

Statement of Problem

As an academic affairs administrator, I have developed an interest in conducting research on college student success as a result of countless conversations and initiatives aimed at improving institutional graduation rates. The problem that my research aimed to address related directly to reducing the attrition rate of STEM majors and the high failure rate in the math and science courses that STEM majors are required to take during their first year in college. Many of the research studies that have examined the causes of student attrition have included students' academic preparation (or lack thereof) as one of the key reasons that students fail to persist (Daley, 2010; Glogowska, Young, & Lockyer, 2007; Johnson, 2006; Tinto, 1993). With regard to STEM, Thompson and Bolin (2011) have cited several major reports on the need for improving our STEM education efforts. As they noted, enrollment in these programs has steadily increased but graduation rates have not. While many colleges and universities have responded by increasing the level of support (e.g., tutoring, supplemental instruction, etc.) they provide for students enrolled in math and science courses, my study aimed to address the problems of high failure rates and the attrition of STEM majors by examining the predictive value of existing cognitive measures of academic preparation.

More specifically, my study involved an analysis of three cohorts of first-year STEM majors at a large, public, urban, Hispanic-serving, research university in the Southeastern U.S. The students who were utilized for my study all entered the institution as freshmen in 2010, 2011, or 2012, and indicated that they planned to major in one of the institution's 15 STEM majors. The students' institutional records were accessed by me to obtain the following information: year they entered college, sex, race/ethnicity, math SAT score, unweighted high school grade point average (GPA), math placement test score, and the highest level of math course completed in high school. In the process of obtaining my sample, I excluded those students who were missing any of these data. I also accessed the university's student records system to gather the final grades that each student received in the first math and science course that he/she took during his/her first year. Students who had not completed math and science courses during the first year or who had enrolled in non-STEM math or science courses were also excluded from my sample.

Once I gathered all of the data for my sample, appropriate statistical analyses were conducted to determine if the traditional cognitive measures of academic preparation (math SAT score, unweighted high school GPA, math placement test score, and highest level of math taken in high school) predicted whether or not the students earned a grade of C or higher in the math and science courses they took during their first year. Each predictor variable was also analyzed separately to determine if it was correlated with success in those math and science courses.

Research Questions

For my study, I generated a list of 10 research questions related to the effectiveness of utilizing the four cognitive variables of academic preparation to predict STEM majors' success in math and science courses. The first five questions related to predicting whether students had passed their first math course and the last five questions related to predicting whether students had passed their first science course. The specific predictor variables that were being tested and the control variables are outlined below in each of the ten research questions.

- Q1 - Do traditional cognitive measures of academic preparation (math SAT scores, unweighted high school GPA, math placement test scores, and highest level of math taken in high school) significantly differentiate between STEM students who pass math and those who fail math when controlling for year of entry, sex, and race/ethnicity?

- Q2 - Does a freshman STEM major's math SAT score account for a significant amount of unique variance when predicting success in the first math course taken when controlling for unweighted high school GPA, math placement test score, highest level of math taken in high school, year of entry, sex, and race/ethnicity?

- Q3 - Does a freshman STEM major's unweighted high school GPA account for a significant amount of unique variance when predicting success in the first math course taken when controlling for math SAT score, math placement test score, highest level of math taken in high school, year of entry, sex, and race/ethnicity?

- Q4 - Does a freshman STEM major's math placement test score account for a significant amount of unique variance when predicting success in the first math course taken when controlling for math SAT score, unweighted high school GPA, highest level of math taken in high school, year of entry, sex, and race/ethnicity?

- Q5 - Does a freshman STEM major's highest level of math taken in high school account for a significant amount of unique variance when predicting success in the first math course taken when controlling for math SAT score, unweighted high school GPA, math placement test score, year of entry, sex, and race/ethnicity?

- Q6 - Do traditional cognitive measures of academic preparation (math SAT scores, unweighted high school GPA, math placement test scores, and highest level of math taken in high school) significantly differentiate between STEM students who pass science and those who fail science when controlling for year of entry, sex, and race/ethnicity?

- Q7 - Does a freshman STEM major's math SAT score account for a significant amount of unique variance when predicting success in the first science course taken when controlling for unweighted high school GPA, math placement test score, highest level of math taken in high school, year of entry, sex, and race/ethnicity?

- Q8 - Does a freshman STEM major's unweighted high school GPA account for a significant amount of unique variance when predicting success in the first science course taken when controlling for math SAT score, math placement test score, highest level of math taken in high school, year of entry, sex, and race/ethnicity?

- Q9 - Does a freshman STEM major's math placement test score account for a significant amount of unique variance when predicting success in the first science course taken when controlling for math SAT score, unweighted high school GPA, highest level of math taken in high school, year of entry, sex, and race/ethnicity?

- Q10 - Does a freshman STEM major's highest level of math taken in high school account for a significant amount of unique variance when predicting success in the first science course taken when controlling for math SAT score, unweighted high school GPA, math placement test score, year of entry, sex, and race/ethnicity?

Assumptions

As in any research study, the researcher is making a number of assumptions in an attempt to address the problem presented. First, despite the fact that the literature on college student success supports the notion my study was guided by an assumption that traditional cognitive measures of academic preparation such as SAT scores, GPA, placement test scores, and courses taken in high school provide the most useful measure of a student's academic capabilities. Along those lines, the study assumed some level of comparability between those measures for all entering students regardless of their previous educational experiences (i.e., the type and location of the schools they have attended). There is a general consensus (despite those who dissent), for example, that SAT scores should be used to compare college applicants since the test is standardized and statistically valid (Burton & Ramist, 2001; Camara & Echternacht, 2000; Fuertes & Sedlacek, 1994; Patterson, Mattern, & Swerdzewski, 2012). Utilizing a college freshman's high school GPA or the highest level of math course he/she took in high school assumed some level of uniformity for those measures regardless of the type of high school attended, where it is located, or the school's level of resources. In other words, the researcher assumed that a student who completed a calculus course at one high school received a comparable level of instruction and knowledge acquisition as did a student who took calculus at a high school across town or in another state or country. According to Cox (2000), colleges and universities often overlook the fact that students' basic knowledge and skills are variable. To avoid making that error, my research study utilized multiple measures of cognitive preparation and academic success in an effort to inform the placement of students in courses that match their academic preparation.

Delimitations

My study was delimited by the fact that it only includes undergraduate STEM majors who entered a particular university as freshmen during one of the following semesters: Summer 2010, Fall 2010, Summer 2011, Fall 2011, Summer 2012, or Fall 2012. All of the students in the sample population attended the same large, public, urban, Hispanic-serving, research university located in the Southeastern part of the U.S. Only students who were majoring in one the institution's 15 undergraduate STEM degree programs and who attempted math and science courses required for STEM majors during their first year were included. The study was also delimited by the use of only traditional cognitive measures of academic preparation as predictor variables such as math SAT scores, unweighted high school GPA, math placement test scores, and the highest level of math taken in high school. The study did not include other demographic factors that have also been shown in the literature to correlate with success in college, such as socio-economic status and family educational background, because those factors are often not tracked by institutions in the same way that cognitive measures are tracked (Daley, 2010). Finally, the study utilized the aforementioned measures to predict success without considering other potentially relevant psychological or psychosocial factors, such as the amount of time spent studying, self-perceived confidence, or level of motivation. Once again, those non-cognitive factors are not widely measured and also rely heavily on students' self-reported perceptions. By only utilizing traditional cognitive measures of academic preparation, my study aimed to provide a predictive model that relied on information that is more readily available to higher education professionals.

Operational Definitions

- STEM – an acronym for Science, Technology, Engineering and Mathematics.
- STEM Majors – the STEM majors at the institution utilized for my study included: Biology, Chemistry, Earth Science, Geoscience, Physics, Computer Science, Information Technology, Biomedical Engineering, Civil Engineering, Computer Engineering, Electrical Engineering, Environmental Engineering, Mechanical Engineering, Mathematics, and Statistics.
- ALEKS score – an acronym for Assessment and Learning in Knowledge Spaces. All incoming students at the institution utilized for my study are required to take this online assessment for the purposes of math placement.
- SAT – the most widely used college entrance examination.
- Unweighted GPA – a grade point average based on the traditional 4.0 scale that does not award additional points for advanced courses.
- STEM math/science courses – used to refer to the key math and science courses required for first-year students majoring in STEM fields. For my study, they included: Intermediate Algebra, College Algebra, Pre-Calculus Algebra, Trigonometry, Pre-Calculus, Calculus I, Calculus II, Statistics, General Biology I, General Chemistry I, General Chemistry II, Physics I with Calculus, and Physics I without Calculus (and a few additional science courses for those students majoring in Computer Science and Information Technology because those students have more flexibility).
- Year of Entry – refers to the year that the student entered college (2010, 2011 or 2012).
- Sex – refers to self-reported biological sex; male or female.

- Race/Ethnicity – refers to a student's self-reported identification with a common group of cultural customs/ancestry. For my study, students were classified as one of the following: Hispanic, Non-Hispanic White, Black/African-American, and Other/Asian American/Not Reported.
- Large, Public, Urban, Hispanic-Serving, Research University – classifications used to describe the institution utilized for my study. Large refers to the fact that the institution enrolls over 20,000 students (the institution, in fact, enrolls approximately 50,000 students). Public refers to the fact that the institution receives funding from the state and federal government and is subject to state reporting and regulations. Urban refers to the fact that the university is located in a metropolitan area and, as such, provides resources and access to its local inhabitants. Hispanic-Serving is a general designation that is given to institutions that have a student population that is at least 25% Hispanic (the institution utilized in my study, in fact, has a Hispanic student population that is consistently higher than 60%). A Research University is a classification for institutions that include a commitment to conducting academic research as part of their institutional mission.

Summary

My research study aimed to contribute to the literature related to college student success, persistence, and retention. In addition, the results of my study provide original research that colleges and universities can utilize to inform efforts to predict whether STEM majors are likely to succeed in required math and science courses. As higher education institutions continue to compete for limited resources, the pressure to produce more college graduates in the STEM fields continues to mount (Thompson & Bolin, 2011). Not only are colleges and universities utilizing this trend to secure additional grants and funding that are being designated for knowledge production in these fields, state and federal governments are also creating initiatives to encourage and reward those institutions who excel in producing STEM graduates.

Recently, the president of the institution utilized for my study participated in a conference on STEM as part of the institution's recognition for granting STEM degrees to large numbers of minority students. This added pressure to enter STEM fields has also trickled down to the students themselves and, as a result, has increased the number of students who enter college with the intent to pursue STEM degrees (Levin & Wyckoff, 1991). What has not necessarily increased, however, is the academic preparation of students entering colleges and universities, especially in the areas of mathematics and science.

As previously mentioned, I conducted my study with the hope of adding to the literature on student success and to those studies that have examined the ability to predict success in STEM. It was also developed to inform institutional policies regarding the support that first-year students need when faced with the rigorous math and science curricula associated with STEM majors. Along those lines, my study utilized four specific cognitive measures of academic preparation to measure their significance for predicting the likelihood of success in those math and science courses. With a greater understanding of first-year STEM majors' readiness for the required math and science courses, colleges and universities can ensure that students enroll in courses that align with their current level of academic skills and provide greater support to those students who are less likely to succeed. This can serve to improve both the experience and academic success of first-year students, which has been shown to be critical to the overall retention and eventual graduation of today's college students (Siegel, 2011; Stovall, 2000).

CHAPTER II
REVIEW OF THE LITERATURE

Over the past several decades, higher education scholars and those interested in college student success have conducted both qualitative and quantitative studies in an effort to assist colleges and universities with developing strategies for supporting that success. For the most part, that research has focused on identifying ways for institutions to improve the poor retention and graduation rates that were highlighted at the beginning of Chapter I. According to Tinto (1993), educators need to continue to increase the number of students who earn college degrees in order to remain competitive in the global knowledge economy. This notion has been reinforced by the current U.S. President, Barack Obama, who has challenged higher education institutions to increase the percentage of citizens with 2-year and 4-year degrees. In order to achieve that goal there has been an increasing emphasis on creating greater access to higher education, which has raised the question of whether today's college students have the academic preparation needed to be successful in college. More recent work by Tinto (2008) has reinforced this concern by highlighting the fact that access to higher education has increased but the percentage of degree completion has not (and has actually decreased for certain populations of students).

Initially, my research study aimed to examine and contribute original work to the literature on college student retention. As the research progressed and I delved further into the existing literature, the study evolved into one that aimed to identify ways to predict college students' capacity or likelihood for achieving early academic success. As a result, this chapter highlights the review of not only the literature on retention, but also the literature on predicting success in college, factors associated with predicting success in STEM fields or their associated mathematics curricula, concerns over using cognitive measures and standardized test scores (like SAT) to predict success, the importance of pre-college preparation, and the rationale for utilizing college GPA and grades to measure success.

College Student Retention

Even though my study focused on predicting student success, the literature review began with examining how the research on retention has provided the framework and foundation for the research related more directly to college student success. Chapter I highlighted several of the retention studies that helped shape the development of the problem and purpose for my research study. Those studies helped reinforce the fact that over the past few decades, colleges and universities have been pressured into utilizing retention and graduation rates as evidence of their effectiveness (Schugurensky, 2003; Watson, 2010). As mentioned previously, there are several studies from the last two decades that have indicated that U.S. colleges and universities are (on average) graduating around 50-56% of their students within six years (Stratton et al., 2008; Freeman et al., 2007). Given the country's ever-changing demographics, the concept of retention is perhaps even more critical for those institutions that provide greater access and educational opportunities for traditionally underrepresented populations (Martin & Meyer, 2010). This was reinforced by Young, Johnson, Hawthorne, and Pugh (2011) who presented data showing how college attendance figures for both African-Americans and Hispanics/Latinos have steadily increased, but that their retention and graduation rates have remained constant. They also found that first-generation college students were more than twice as likely as their non-first-generation counterparts to leave after one year. Smith (1995) also provided data to draw attention to the lower retention rates of underrepresented minority students, noting that this was an even more critical problem for those students pursuing STEM degrees.

One of the challenges with the literature on retention has been that it has focused primarily on what institutions can and should be doing to influence retention and inform retention initiatives without considering the characteristics of the students themselves. Otero et al. (2007), and Williford and Wadley (2008) illustrated that by highlighting the fact that most retention studies have focused on what institutions can do to improve retention and often ignore the motives or factors related to why students decided to leave and what happens to those who do choose to leave. While examining an institution's retention strategies is definitely a key piece of solving the retention puzzle, educators must also consider how those strategies address students' academic preparation. According to Belchier, Michener, and Gray (1998), we need to accept the reality that the decisions linked to leaving college are as diverse as the students themselves. As mentioned in the last chapter, more than half of the variance in institutional retention rates is directly related to attributes of the students rather than institutional factors (Freeman et al., 2007). Those student attributes include the student's academic profile and the capacity for completing college-level work. That is why experts on college retention agree that a major obstacle to improving retention involves addressing the fact that many students are academically unprepared for the rigors of college (Daley, 2010; Stratton et al., 2008; Tinto, 1993).

The notion of academic preparedness has also served as a major motivation for establishing admission standards as a means for predicting student success, especially during times of increased access to higher education. In order to establish those standards institutions should look to the literature for guidance. The retention literature contains several examples of studies that have described the use of qualitative measures to identify those factors that might help an institution predict attrition or that might influence a student's decision to stay or leave (Glogowska et al., 2007; Lehmann, 2007; Otero et al., 2007; Williford & Wadley, 2008; Woosley, Slabaugh, Sadler, & Mason, 2005). Other studies have utilized quantitative methods to analyze the obstacles that might prevent college students from persisting and the reasons they might consider leaving college, as opposed to analyzing students who had already left (Johnson, 2006; Freeman et al., 2007). In each of those studies, one of the consistent factors that emerged was students' perceptions that they had academic deficiencies or that they were having difficulties completing college-level work. For example, Freeman et al. (2007) utilized their College Student Attrition Survey (that they developed specifically for their research) to identify the most common variables associated with students who had considered dropping out of college. Along with psychological variables such as “social life” and “lack of diversity,” they found that one of the most significant variables for identifying students who were at-risk of leaving was their perceptions of being academically unprepared for college. Along those same lines, Johnson (2006) noted that college student attrition is strongly associated with poor college grades and below-average academic performance. Daley’s (2010) research also highlighted the fact that students’ lack of self-knowledge regarding their academic preparation contributed to their lack of success and potential decision to depart. The ways that this combination of factors (students who are academically unprepared, poor college grades/performance, and lack of self-knowledge regarding academic abilities) impact college student retention suggests that we need to do more to predict whether or not students can be successful, especially in the courses in which students traditionally struggle most.

The research on retention has also highlighted the fact that there are several non-cognitive factors that influence a college student’s decision to persist. As Mathiasen (1984) pointed out, if retention was merely linked to academic potential then every student who did well on the SAT would ultimately succeed in college (when, in fact, there are many who do not). Wheat, Tunnell, and Munday (1991) found that student success can depend on attitude and several factors besides just aptitude. For example, the work by Belchier et al. (1998), Glogowska et al. (2007), and Blanchard and Mascetti (2000) all found that even the best, most effective retention strategies will not work with students who decide that they do not want to be there. Kanoy, Wester, and Latta (1989) pointed out that non-traditional predictors such as locus of control, psychological variables, and academic self-concept are also important for understanding student attrition. Stratton et al. (2008) addressed this by noting that “dropout behavior is explained as a rational response to new information that changes the probability with which one will receive a degree and/or the costs/benefits associated with that degree” (p. 320). Cole and

Espinoza’s (2008) research supported Stratton’s perspective on attrition and added that factors such as connection to campus and ability to foster relationships with faculty have also contributed to student persistence and retention. The work that Burton and Ramist (2001) produced for the College Board provided an even more thorough assessment by suggesting that institutions use these non-academic measures in conjunction with the more traditional cognitive measures to improve the validity of establishing admission criteria that align with predicting student success.

With regard to the retention studies that focus on predicting academic success, there is a nascent body of literature that has begun to consider students’ pre-college preparation. That research utilizes those pre-college characteristics in an effort to evaluate the factors that can best predict success and persistence (Engle & Tinto, 2008; Johnson, 2006; Otero et al., 2007; Tai, Ward, & Sadler, 2006). Johnson’s (2006) work, in particular, has referenced Tinto’s Student Integration Model and Bean’s Student Attrition Model to highlight the fact that “students’ decision to persist is determined by the quality of ongoing interactions between pre-college characteristics and institutional environments” (p. 907). According to Engle and Tinto (2008), a rigorous high school curriculum, including advanced mathematics courses, increases the likelihood that students (especially low-income and first-generation students) will attend and succeed in college. The research studies referenced in this paragraph were actually responsible for shifting my research study away from purely understanding retention patterns and toward examining the traditional cognitive measures of students’ pre-college preparation to try and predict first-year success. Along those lines, the next section of this chapter will provide a review of the literature related to predicting students who are likely to succeed in college.

Predicting Success in College

Generally speaking, the research that has been conducted on predicting student success has focused on helping colleges and universities with admission decisions. The majority of that research has also focused on using traditional cognitive measures, especially SAT scores and high school GPA, as the primary tool for determining which students have the greatest potential to succeed. As mentioned in the previous section, there is a growing body of research that looks beyond traditional cognitive measures and examines the ways that more psychosocial factors such as motivation, attitude, and commitment to earning a degree impact student persistence. My research study, however, focused on the more cognitive measures that are being utilized to predict success. According to Fuertes and Sedlacek (1994), academically-related variables are the best predictor of a student's future grades. On an even more general note, Levin and Wyckoff (1991) framed their research utilizing the notion that the best predictor of future behavior is past behavior. While they did include some psychological measures in their study, they emphasized the use of several academic and cognitive factors to predict the success of engineering majors at Pennsylvania State University (Penn State). During the review of the literature, I found that most studies on predicting college success still viewed the cognitive and academic factors as providing the most predictive power. Kanoy et al. (1989), for example, conducted a study that involved developing a model that utilized both cognitive and psychological measures to predict the success of college freshman. While they found that academic self-concept (a psychological factor) to be an important predictor for certain students, the traditional cognitive predictors (especially high school GPA) accounted for the majority of the variance in predicting college GPA.

For decades, colleges and universities have been utilizing standardized tests like the SAT as an admissions examination to assist with determining if applicants have the potential to succeed in college. That practice, while increasingly controversial, has been supported by several researchers who have sought to validate its predictive utility. Studies by both Fuertes and Sedlacek (1994) and Zwick and Sklar (2005), for instance, highlighted several bodies of research that have found the SAT to be effective in predicting first-year college grades. They did note, however, that the SAT has not proven sufficient to predict college success/GPA beyond the first year. The work conducted by Burton and Ramist (2001) utilized over 15 years of SAT and college GPA data to confirm that the standardized test scores did, in fact, help predict which students excelled academically in college. Since the institution utilized for my study enrolls a significant minority population, the literature review also examined what researchers had to say about the use of SAT scores to predict the success of traditionally underrepresented populations. While there are some who might challenge the results, each of the research studies mentioned above found that the SAT was a valid predictor for all students, including the ethnic minorities in their samples. Fuertes and Sedlacek (1994), in particular, noted that the SAT scores they analyzed were positively correlated with the success of the Hispanic students that were included in their study. Based on that review of the relevant research, SAT scores (the math subsection) were included as one of the predictor variables for my research study. The only case against using standardized scores that surfaced was related to the SAT's lack of long-term predictive power. However, since my study focused on first-year students, that critique was not a major concern.

In addition to the SAT, several researchers have also examined the practice of utilizing a student's high school performance as a predictor for college success. That has typically involved utilizing a student's high school GPA or his/her class rank. Overall, this strategy assumes a belief in the philosophy (mentioned in the beginning of this section) that educators should rely on past behavior to predict future behavior. Several research studies that were reviewed have found that high school GPA is typically the single best indicator of college success (Chase & Jacobs, 1989; Hoffman & Lowitzki, 2005; Williford, 2009; Zwick & Sklar, 2005). Several of those studies have actually compared the predictive power of the high school GPA with that of the SAT and other cognitive and psychological factors. Zwick and Sklar (2005) reviewed several research studies, including those that had been conducted by the ACT organization, the Association for Institutional Research, the College Board, the Educational Testing Service (ETS), and the National Association for College Admission Counseling, all of which found that high school

GPA or class rank was the most important factor in predicting future academic persistence (the SAT was consistently the second best predictor). A recent research study conducted by Columbia University also concluded that high school grades are better predictors of success than standardized tests (Belfield & Crosta, 2012). Thompson and Bolin's (2011) research concurred and also found that a significant relationship existed between a student's likelihood to drop out of college or change his/her major and high school rank. With regard to utilizing a student's high school GPA, institutions and researchers alike also have begun to examine whether a student's weighted GPA (which includes additional points for advanced courses) or unweighted GPA (which is based on the standard 4.0 scale) will provide more predictive power. According to Nagaishi and Slade (2012), who analyzed the high school transcripts of over 500 pre-med students in Texas, unweighted GPAs were more useful for predicting those students' academic success in college. Each of those research studies, especially the one that linked high school grades to selection of major, confirmed my decision to include high school unweighted GPA as part of the analysis being conducted for my study (note: the institution utilized for my study does not record its students' high school class rank).

Given that both SAT scores and high school GPA have been found to be useful in predicting students' college GPAs, several of the research studies also addressed the prospect of utilizing a combination of both of those measures. As Noble and Sawyer (2004) pointed out, colleges and universities typically use both high school grades and test scores (like SAT) to predict their applicants' probability for success. In addition to assisting institutions with improving their academic reputation and ranking, this common admission practice can be linked to what many educational researchers have offered as a tool for predicting students' academic potential. The research conducted by Mathiasen (1984), for instance, reviewed over 60 studies that all confirmed that high school academic performance and admission test scores (SAT/ACT) are the best predictors of college success. Camara and Echternacht (2000) also reviewed several studies and concluded that both SAT scores and high school GPA were highly correlated with various measures of student success. The research conducted by Harackiewicz, Barron, Tauer, and Elliot (2002) evaluated the use of these two cognitive measures against several others, including psychological and psychosocial factors. They found that both student ability (as measured by SAT scores) and prior high school performance (as measured by their high school GPA or rank) contributed a significant amount of unique variance in predicting college academic performance. While all of this was helpful for affirming my study's research variables, the most validating point related to those past studies that utilized a variable that combined the high school GPA with the SAT score. In each of those studies, the researchers found that high school GPA was a better single predictor of success, but that the combined variable (GPA and SAT) was consistently better than using either one or the other (Burton & Ramist, 2001; Chase & Jacobs, 1989; Hoffman & Lowitzki, 2005).

In addition to the standardized testing that institutions often use for admission, the literature also included recommendations regarding the use of placement tests that new students are often required to take prior to enrolling in college. According to Wheat et al. (1991), the use of placement tests is desirable because institutions cannot always trust the subjective nature of high school grading policies. Cox (2000) supported that recommendation as part of his study that aimed to identify the knowledge and skills required for success in both English and mathematics courses. He suggested that the use of diagnostic (placement) tests was necessary since institutions should not assume that students have acquired a certain level of knowledge based solely on their high school grades. Along those lines, Scott-Clayton (2012) cited a study that found that placement tests were valid for predicting success in college-level mathematics courses. On an even more promising note, the research conducted by Veenstra, Dey, and Herrin (2008) uncovered a study that found that students' math placement test results were a significant factor in predicting their first-year GPA (overall, not just in math). While it would be interesting to discover if those results could be replicated in future studies, my research study aimed to consider the sample population's math placement test scores in conjunction with other factors. That strategy was supported by the work conducted by Armstrong (2000), who found that the overall validity of placement test scores was weak but that their predictive value increased when colleges in California combined them with other academic measures such as high school GPA.

As I have just described, there have been several research studies that have attempted to identify the factors that can best predict academic success in college. While many of those have included non-cognitive measures, the choice of predictor variables for my study was justified by the fact that traditional cognitive measures are still the most common and most reliable tools for predicting academic success. Beyond that, what this portion of the literature review helped confirm was the fact that there are multiple variables that provide significant predictive value. According to the successful model developed by Kanoy et al. (1989), colleges and universities need to use multiple predictors to more accurately determine a student's academic potential.

Predicting Success in STEM and Mathematics

While the majority of the student success literature has focused on predicting college success in general, there have been some empirical studies at institutions that aimed to predict the success of students in certain STEM fields or in the rigorous mathematics courses that are required by those disciplines. The research suggests that the factors that best predict success in STEM are basically the same factors that predict students' overall success. Standardized test scores (both SAT and placement tests), high school grades, and courses taken in high school are all significant factors in predicting STEM success. As mentioned in Chapter I, there have been several reports concerning the need to improve STEM education (Thompson & Bolin, 2011). With that in mind, predicting the success of students pursuing STEM majors would seem to be a vital aspect of those efforts. According to Kessel and Linn (1996), many students who enter college with the intent to pursue a math or science major end up changing their minds. As they noted, as many as two-thirds of those students who begin as math or science majors eventually graduate with degrees outside of those disciplines.

A recent study that was conducted by Thompson and Bolin (2011) examined the retention and graduation rates of students attending a large, public university in Texas. More specifically, they compared the students pursuing STEM majors to those who were studying business and education. Not only did they find that the STEM majors were more likely to change their major, they also found that the STEM majors were overall less likely to graduate than their peers in the business and education disciplines. That was particularly disturbing to the researchers given that those students in STEM majors were not any less prepared academically (and in some cases had even higher academic credentials) than the non-STEM students. With regard to those students who change majors, Haislett and Hafer (1990) have noted that while many students choose to leave STEM fields because of a change in career goals, many others do so because of academic difficulties. It seems logical to them, then, that we work to devise a method for predicting who might consider leaving the STEM disciplines for academic reasons.

According to Levin and Wyckoff (1991), as enrollment in the engineering fields has increased, so has attrition. The study that they conducted with Penn State engineering students utilized 19 variables (including traditional cognitive measures) to identify the factors that were most linked to high college GPAs. As a result, they found that high school GPAs above 3.0, math SAT scores of 600 or higher, and placing directly into Calculus I (based on placement test scores) were the factors that correlated most positively with the college GPAs of engineering majors. A study conducted by Sadler et al. (2012) found that success in high school math and science courses (especially Calculus I) were strong predictors of success in engineering programs. Pressures to improve the retention of engineering students at the University of Michigan prompted Veenstra et al. (2008) to develop a study to try and understand the characteristics linked with student success. Not only did they find that both the math SAT score and high school GPA were significant predictors of success, they also noted that 38% of the variation in engineering students' first-year GPA was attributed to their pre-college factors, especially mathematics preparation.

According to Cole and Espinosa (2008), high school performance and academic preparation are highly correlated with college success for not just engineering majors but for STEM majors in general. They also found that to be true for the minority students who were included in their study, which was of particular interest given the significant number of STEM degrees that the institution utilized for my study has awarded to minority students. A group of researchers from the University of Pittsburgh and the University of Alabama in Huntsville have also conducted multiple studies in an effort to identify variables that might predict both potential interest in and aptitude for STEM disciplines (Nicholls et al., 2007; Nicholls, Wolfe, Besterfield-Sacre, & Shuman, 2010). They found that both math SAT scores and high school GPA were significant indicators of students who had intended to major in STEM fields. Their research also focused on the importance of math in predicting the capacity for success in STEM and noted that students who were struggling to keep up with the math curriculum by eighth grade were less likely to succeed in STEM disciplines. In their studies, as well as the previous studies that they reviewed, math aptitude and ability was significantly more important for those students who chose to pursue STEM degrees. That was also the case for the research conducted by Tai et al. (2006) who examined the variables related to succeeding in college chemistry. In addition to students' previous exposure to chemistry, math SAT score and calculus enrollment (in high school) were highly significant in predicting grades earned in their first college chemistry course.

Given the critical nature of mathematics for those in the STEM disciplines, the literature review also included the research related to predicting success in college-level math courses. A study involving college students in Australia concluded that both prior math achievement and attitude toward math were highly predictive of later success (Hemmings, Grootenboer, & Kay, 2011). According to Kessel and Linn (1996), both previous grades and entrance exams (i.e., SAT) significantly predict success in college math courses. While they warned against using math SAT scores as the only predictor, they did note that those scores are useful in predicting success when combined with math placement test results. With regard to the benefits of using placement tests for predicting math success, studies by Wheat et al. (1991) and Cox (2000) confirmed that placement testing is both a common tool and highly correlated with students' actual success in college math. The study by Wheat et al. (1991) actually found that high school grades and placement test results were the best predictors of students' ability to succeed in College Algebra (a common course required by many STEM disciplines). Given the fact that many STEM students have decided to change their majors due to academic difficulties, this emphasis on predicting success in math supports the need for ensuring that a student's first-year curriculum aligns with an academic level (especially in math) that matches his/her abilities. The research presented in this section supported my decision to utilize factors such as high school GPA, math SAT scores, math placement test scores, and previous math courses to try and predict the success of students pursuing STEM degrees.

Concerns Related to Using Cognitive Measures

As mentioned previously, many researchers have begun to include non-cognitive factors to paint a broader picture of students' academic potential. The relevant literature also included a number of studies and commentaries that have challenged the validity of using the traditional cognitive measures to predict who will or will not be successful in college. Zwick and Sklar (2005) noted that bodies of research have emerged with the sole purpose of determining the effectiveness of utilizing both high school GPA and SAT as predictors of college success. According to Kanoy et al. (1989), there have been many educational researchers who have challenged the use of these traditional cognitive factors, especially with regard to their ability to predict the success of minority students. More specifically, those critiques have been directed toward the use of standardized tests such as the SAT and ACT as the only measure for predicting success.

According to Fuertes and Sedlacek (1994), the SAT was designed "to assess the scholastic ability of high school students entering college" (p. 350). Harackiewicz et al. (2002) reinforced that by describing the ways that colleges and universities have utilized SAT scores to quantify students' abilities and academic potential. Despite the fact that it is a common practice to include SAT scores in admission decisions, there has been a great deal of controversy surrounding their effectiveness in predicting future success. Back in 1996, Kessel and Linn pointed out that there have been several studies that have argued against using only SAT scores because they have been shown to underpredict college grades and overall academic success. Cimetta, D'Agostino, and Levin (2010) also cited several studies that have challenged the use of SAT scores as part of their research on the prospect of using state high school achievement tests (in Arizona) as a substitute for SAT scores. One of the biggest complaints regarding their use has been the claim that SAT questions are racially biased. Hoffman and Lowitzki (2005), for example, claimed that this bias makes the use of standardized tests ineffective for predicting the success of certain minority populations.

These critiques of using the SAT to predict success were a source of concern related to the decision to include SAT scores as one of the predictor variables for my study. There were, however, multiple research studies that reinforced the validity of including SAT scores. As mentioned previously (and which will be discussed further in Chapter III), the work of Burton and Ramist (2001) and the research conducted by Camara and Echternacht (2000) have both confirmed the validity of using both SAT scores and high school GPA for predicting the academic success of first-year college students. In addition, the recent research from Patterson et al. (2012) supported the predictive validity of utilizing SAT scores. As they pointed out, there have been several studies that have shown that a student's highest SAT score (for those who have taken the test more than once) correlates highly with his/her first-year college GPA. Since my study involves predicting the success of first-year STEM majors, math SAT scores were ultimately included as one of the predictor variables.

Importance of Pre-College Preparation

Throughout my review of the pertinent literature, the relationship between college performance and high school preparation was referenced repeatedly. In fact, the research has shown that a significant amount of variation in college GPAs can be explained by pre-college preparation and academic performance in high school (Sadler & Tai, 2001; Williford, 2009). According to Cole (2001), the human brain typically uses past experience to accept and reject information that might conflict with what one has come to know as true. Perhaps that is why many college students struggle when they attempt to approach college utilizing only the tools they acquired in high school. Highlighting the concerns over the lack of academic rigor in high school, researchers have noted that colleges often blame high failure rates on the perception that high school courses have not equipped students with the skills needed to handle college-level work (Hoyt & Sorensen, 2001; Roth, Crans, Carter, Ariet, & Resnick, 2001). That has led many educators to suggest that increasing the academic rigor in high school will not only help students with college courses but will also provide them with more confidence in their academic abilities. As Jalomo (2000) pointed out, students who develop an academic self-concept early on are less likely to experience academic difficulty and more likely to graduate from college.

Along the lines of increasing academic rigor, Williford (2009) noted that several studies on improving college student success (especially in the first year) have recommended that students take more rigorous courses while in high school. In fact, the studies that he reviewed found that the intensity of the high school curriculum mattered more than any other pre-collegiate factor in predicting student success. As mentioned previously, the research conducted by Veenstra et al. (2008) also emphasized the importance of high school preparation in predicting the first-year GPAs of students majoring in engineering. That notion of academic preparation has been found to be significant for other STEM majors and for students from underrepresented populations as well. According to Cole and Espinosa (2008), the skill development and academic performance that students achieved prior to college served as the best indicators of success for minorities pursuing science-related majors. Given the traditionally challenging curriculum for STEM degree programs, it seems logical that students pursuing STEM majors would benefit from more exposure to rigorous mathematics and science courses while they are still in high school. According to Burton and Ramist (2001), the most stringently graded college subjects are science, engineering, and calculus. Since mathematics is the common denominator between those subjects, evaluating students' math preparation has become a popular theme in educational research.

As reported by Hoyt and Sorensen (2001), college professors have lamented that their students have not acquired the math skills in high school that are necessary for succeeding in college. According to Cox (2000), the level of math preparation that a student obtains prior to college has been shown to be important to predicting his/her probability of succeeding in college-level math courses. For many researchers, the key to increasing that level of math preparation involves the amount and type of math courses that students take prior to college. The research being done to predict the likelihood for success in STEM has shown that by eighth grade we can already ascertain a student's capacity to succeed by evaluating his/her math abilities (Nicholls et al., 2010). Levin and Wyckoff's (1991) research with Penn State engineering students found that those students who had enough high school math preparation to enable them to place into and start with Calculus I (in college) were more likely to persist and graduate with an engineering degree. According to Sadler and Tai (2001) and Tai et al. (2006), students who had taken a calculus course in high school consistently performed better in college physics and chemistry (which are both courses that are required by several STEM disciplines).

This aspect of the literature review was particularly influential in my decision to include the highest level of math taken in high school as one of the predictor variables for my research study. According to Davis and Shih's (2008) research, the number of years of high school math that students completed had a statistically significant influence on both their math placement and their first math grades in college. That is why my study examined the math courses that the students took in high school to determine if a relationship existed between that variable and their ability to pass their first math and science courses in college. My review of the literature also revealed that utilizing students' performance in individual high school courses (like math) to predict future success has not been studied sufficiently. For that reason, the results of my study will contribute original research to the literature on predicting college student success. The literature on high school preparation also reinforced the decision to utilize multiple predictors to predict success since, as Burton and Ramist (2001) found, college students' grades in the rigorous science, engineering, and calculus courses have a relatively low correlation with any one cognitive predictor.

Rationale for Utilizing GPA/Grades to Measure Success

As the research design for my study was developed, an intentional effort was made to review the literature that would inform how to define the dependent variables and measure the sample population's success. According to both Noble and Sawyer (2004) and Zwick and Sklar (2005), the majority of research that aims to predict college success utilizes students' first-year GPA. In fact, most of the research studies that I reviewed emphasized the use of either first-year GPA or first-year grades as the predominant measure of college student success (Burton &

Ramist, 2001; Chase & Jacobs, 1989; Fuertes & Sedlacek, 1994; Johnson, 2006; Noble & Sawyer, 2004; Patterson et al., 2012; Veenstra et al., 2008; Zwick & Sklar, 2005). The work of Thompson and Bolin (2011), who have examined critical efforts to improve STEM education, reinforced the importance of the first year by noting that the first year is when the highest number of dropouts has occurred.

While utilizing first-year GPA as the dependent variable was something that was also considered for my study, I ultimately decided to focus more on the students' success in the key math and science courses that STEM majors are required to take during their first year. According to Burton and Ramist (2001), the validity literature supports very few alternatives to utilizing GPA as the measure of success. However, the work of Camara and Echternacht (2001) proved useful in confirming that both SAT scores and high school GPA are valid for predicting multiple criteria for success, including GPA, graduation rates, and course grades. Johnson's (2006) research also reaffirmed the benefits of utilizing course grades as a measure of student success. As he noted, most empirical studies have shown that grade performance (as opposed to overall GPA) at the end of the first term is the most important factor in predicting college student persistence. Along those lines, Sadler and Tai (2001) and Tai et al. (2006) utilized course grades in their students' first-year physics and chemistry courses (respectively) as their measures for differentiating between successful and unsuccessful students. With that in mind, it seemed justifiable to utilize STEM majors' grade performance in the math and science courses they take during their first term as an early indication of whether or not they have the potential to succeed in and graduate with STEM degrees.

Summary

Conducting a review of the pertinent literature provided an array of useful information related to the efforts to predict first-year STEM majors' likelihood of passing their first math and science courses. More specifically, that review affirmed the usefulness of the predictor variables that were selected for my study. As was previously mentioned, despite the concerns that have been raised over utilizing cognitive variables to predict college student success, there are several research studies that support the validity of utilizing both SAT scores and high school GPA (Burton & Ramist, 2001; Camara & Echternacht, 2000; Patterson et al., 2012). The work of Burton and Ramist (2001), in particular, highlighted the fact that the research studies that have included both of those measures of cognitive ability have been consistently better than studies that have only used one or the other. Along those same lines, Zwick and Sklar (2005) pointed out that even though GPA and SAT are the best predictors, when only one of them is used there is a lower correlation in predicting first-year GPA for both African-American and Hispanic/Latino students. This concern over relying on one (or even two) variables to predict student success was also echoed by the research of Armstrong (2000), Cole and Espinosa (2008) and Kanoy et al. (1989), all of whom supported the benefits of utilizing multiple predictors.

Looking beyond SAT and GPA, several of the studies presented in this chapter reinforced the use of math placement test scores as a measure for both improving and predicting success in college math courses (Scott-Clayton, 2012; Cox, 1998; Wheat et al., 1991). According to Williford (2009), there have not been many studies that have looked at specific high school courses to determine their usefulness in predicting college success. During my review of the literature, however, I found studies that supported the fact that evaluating a student’s math preparation (based on the courses that he/she has completed in high school) can significantly predict his/her placement and success in college math and science (Camara & Echternacht, 2000; Sadler & Tai, 2001; Tai et al., 2006; Wheat et al., 1991). With regard to measuring student success, most of the research has utilized first-year GPA. In support of my decision to measure success by analyzing whether or not students passed specific math and science courses, I relied on the research of Camara and Echternacht (2000), Sadler and Tai (2001) and Tai, Ward, and Sadler

(2006) as support for utilizing traditional cognitive measures to predict success in individual courses. Finally, while the literature review did reveal several studies related to predicting college student success and the success of STEM majors, my study provides an original contribution to the literature. By utilizing multiple cognitive variables (more specifically, four variables that have been utilized within the same study) to predict success in specific math and science courses, my research study revealed new information that can be utilized to assist colleges and universities with improving the experiences and retention of students pursuing STEM degrees.

CHAPTER III
RESEARCH DESIGN

As outlined in Chapter I, my study involved an analysis of three cohorts of first-year STEM majors at a large, public, urban, Hispanic-serving, research university in the Southeastern U.S. The institution's student records were utilized to analyze the relationship between traditional cognitive measures of academic preparation and students' success in first year math and science courses. Those traditional cognitive measures of academic preparation included students' math SAT score, their unweighted high school grade point average (GPA), their math placement test score (ALEKS), and their highest level of math completed in high school. The year that the students entered college, their sex, and their race/ethnicity were utilized as control variables for the quantitative analyses that were run to address the study's research questions.

In this chapter, I have outlined the overall research design that was utilized to answer the 10 research questions that were provided in Chapter I (on pages 7-9). Generally speaking, the aim of my study focused on my attempt to examine whether specific cognitive measures of academic preparation helped predict whether STEM majors passed the math and science classes they took during their first year in college. Each of the four cognitive measures that were utilized as predictor variables for my study (math SAT score, unweighted high school GPA, ALEKS placement test score, and highest level of math completed in high school) were also analyzed separately to determine if they provided a significant amount of unique variance when predicting success in the math and science courses, while controlling for alternative hypotheses. In addition to the overall statistical design, this chapter also contains more detailed information about how the students were identified and selected, the data collection procedures, and the research and statistical hypotheses that were tested utilizing quantitative methodology.

Sample Population

The students for my study were first-year U.S. college students who entered the same large, public, urban, Hispanic-serving, research university in the Southeastern U.S. during the summer or fall of 2010, 2011, or 2012. Those students had all identified that they planned to major in one of the 15 STEM disciplines offered by the university. In other words, the sample population was drawn exclusively from students who were pursuing STEM majors and who entered college during those three years. Despite the fact that some of the students may have entered the university with college credits that they had earned while still in high school, they were all classified as “first-time in college” students (the designation used by their university for students entering as freshmen, regardless of the number of college credits earned prior to enrolling in college).

The only demographic information that was gathered for the students was their self-reported sex and self-reported race/ethnicity, which are provided in Table 1 (below) and again in Chapter IV. Based on the institution’s overall demographics, however, we can also ascertain that the vast majority of the students were traditional-aged college students (between the ages of 17 and 24) who attended high school in the state where the institution is located. The average age of first-year students is not one of the institution’s reported statistics, but data obtained from the students who attended the mandatory freshman orientation sessions confirmed that nearly all of the students who are admitted as freshmen are 24 years old or younger. In addition, approximately 90% of the institution’s undergraduate students are residents of the state in which the institution is located. Table 1 (below) has been provided to offer a comparison between the sex and race/ethnicity of the sample population and the institution’s overall demographics. With regard to sex, it is worth noting that while the institution’s enrollment is about 55% women and 45% men, the sample was almost 80% men. This discrepancy was expected based on the fact that my study utilized only STEM majors, and men are still far more likely than women to pursue careers in the STEM disciplines in the U.S. (Nicholls et al., 2007; Thompson & Bolin, 2011). With regard to race/ethnicity, the table below provides the compiled data for all first-year students who entered the university during the three years from which the sample was extracted. On average, during those three years, 67.9% of the institution's first-year students were Hispanic, 11% were White/Non-Hispanic, 11.2% were Black/African-American, and 9.9% were classified as "Other." As noted in Table 1, the racial/ethnic breakdown of the sample was almost identical to that of the institution's overall first-year student population.

TABLE 1

Comparison of Institutional and Sample Population Demographics

	Institution	Sample Population
Female	55%	20.9%
Male	45%	79.1%
Hispanic	67.9%	66.9%
White/Non-Hispanic	11%	11.1%
Black/African-American	11.2%	11%
Other	9.9%	11%

Data Collection Procedures

In order to collect the data that was utilized in my study, I retrieved information from the university's student academic information system, which houses all of the academic records for each student. The first step in the data collection involved identifying those students who enrolled as first-time students in college (i.e., freshmen) for the following semesters: Summer 2010, Fall 2010, Summer 2011, Fall 2011, Summer 2012, and Fall 2012. It is worth noting that these particular semesters were utilized because the majority of the freshmen at the institution utilized for my study start in either summer or fall semester each year. Those students admitted for summer semesters are those who typically have lower SAT scores and/or lower high school GPAs but are otherwise deemed to be "college ready." The institution also admits a (much) smaller number of freshmen in spring semesters, but those students were not included as part of the sample for my study. The decision to include only those students who began at the institution in summer or fall was based on the fact that students who enter as freshmen in the spring semester have often completed college courses at a different post-secondary institution during the fall. By excluding those who began in spring, the study sample was limited to those students who were actually enrolled in their first semester of college.

Once the population of students who entered the institution during the designated semesters was identified, the sample was then limited to those students who indicated that they planned to major in one of the 15 identified STEM majors that are offered by the university. Those majors included: Biology, Chemistry, Earth Science, Geoscience, Physics, Computer Science, Information Technology, Biomedical Engineering, Civil Engineering, Computer Engineering, Electrical Engineering, Environmental Engineering, Mechanical Engineering, Mathematics, and Statistics. This sample of students, which totaled 1367 students, was then used to query the student information system and obtain demographic information (sex, race/ethnicity) as discussed in the previous section. In addition, that query extracted the data that were utilized as the four independent/predictor variables for my study. Once again, those four variables were math SAT score, unweighted high school GPA, institutional math placement test (ALEKS) score, and the math courses taken in high school for each of the STEM majors who entered the institution from 2010 to 2012. At that point in the data collection, 145 students were eliminated from the sample because they did not have an SAT score (each of these students had been admitted to the university based on their ACT score, which is the other popular standardized test utilized for college admission). An additional 187 students were eliminated because they did not have an ALEKS score (almost all of these students were from the cohort admitted in 2010 because even though the institution was already utilizing the placement test, the ALEKS scores were not consistently entered into the student records system during that year). For the 1035 students that remained in the sample, their math SAT scores, unweighted high school GPA, and ALEKS scores were recorded in the data set as continuous variables. With regard to the math courses taken in high school, each student's record was reviewed to determine the highest level of math that he/she had attempted. Each student's highest level of high school math was then coded and recorded as an ordinal variable utilizing the methodology outlined in Table 2.

TABLE 2

Method for Coding Students' Highest Level of High School Math

Highest Level of Math Taken	Coded As
Algebra II (or lower)	0
Pre-Calculus/Trigonometry/Analytic Geometry	1
Statistics (for students majoring in Computer Science/IT)	1
Calculus I/Calculus AB	2
Calculus II/Calculus BC	3

Once those 1035 students were identified, the student information system was once again utilized to obtain the grades that each student earned in their first math and science course completed at the university. At that point, an additional 17 students were removed from the sample because they either had no enrollment on record or they had not completed a STEM-related math or science course during their first year at the institution. That information was then utilized to create two binary dependent variables to represent students' success in those math and science courses. As such, students who earned a grade of C or higher in their first math course were assigned a score of "1" and those who earned a grade of C- or lower (or who did not complete the course) were assigned a score of "0." Likewise, students who earned a grade of C or higher in their first science course were assigned a score of "1" and those who earned a grade of C- or lower (or who did not complete the course) were assigned a score of "0." Those scores, based on the grades earned in the students' first math and science courses, were utilized as the nominal dichotomous dependent variables for my study. More specifically, the dichotomous variable related to math grades (1 = passed, 0 = did not pass) was utilized as the dependent variable for research questions 1-5 and the dichotomous variable related to science grades was utilized as the dependent variable for research questions 6-10 (once again, the research questions can be found on pages 7-9 in Chapter I).

Once the students and their corresponding data were identified, those data (math SAT score, unweighted high school GPA, ALEKS math placement test score, coded variable related to highest level of math taken, year of college entry (1=2010, 2=2011, 3=2012), sex (0=female, 1=male), race/ethnicity (1=Hispanic/Latino, 2-White/Non-Hispanic, 3= Black/African-American, 4=Other (Unknown, Not Reported, and Asian/Asian American)), and dichotomous variables related to passing/not passing math and science courses) were entered into SPSS for data analysis purposes. Once the descriptive statistics (i.e., frequencies) were obtained, the three control variables (year of entry, sex, and race/ethnicity) were binary coded in order to determine the significance of each category within those variables.

Research Design

My study was conducted utilizing an ex-post facto research design with and controlling for alternative hypotheses. An ex-post facto design was most appropriate because my study aimed to look for relationships between something that had already occurred and the factors that might have helped predict the outcome (McNeil, Newman, & Fraas, 2012). Since the students had already completed the math and science courses that were analyzed, there was no opportunity to manipulate the variables or conduct any type of experimental research. The hope is that the results of this ex-post facto study will inform future experimental research or institutional practices in which students are placed in different levels of math and science courses based on the predictive power of the independent variables that were analyzed.

The quantitative analysis utilized hierarchical logistic regression to determine how well the four cognitive measures of students' academic preparation significantly differentiated students who passed and did not pass their first math and science courses. Generally speaking, a design utilizing regression was selected because the study attempted to find a correlation between a set of predictor and dependent variables. More specifically, a logistic regression was appropriate since the dependent variables (passing or not passing math and science courses) were nominal dichotomous/binary variables as opposed to continuous variables. The statistics examined were the -2*Log Likelihood statistic, full model Chi-Square, pseudo R^2 (the Cox & Snell R^2 was utilized for my study), and the classification model. The -2*Log Likelihood statistic was utilized because it provided a measure of the unexplained variability in the data. A decrease in the value of that statistic (from block one to block two), provided an indication of how much new variance was explained by the predictor variable(s) being added to the regression model. Statistically speaking, as the model becomes better the value of the -2*Log Likelihood statistic will decrease in magnitude (Norusis, 1998). For each analysis, the Chi-Square statistic and corresponding *p*-value provided an indication of whether the amount of variability being accounted for in the final regression model was statistically significant. The pseudo R^2 statistic was utilized to measure the proportion of variance accounted for by each model (first excluding the variables being tested and then including those variables). The classification model provided information on how the variables being tested influenced the ability to correctly predict the outcome of the dependent variables. The predictor variables included math SAT scores, unweighted high school GPA values, ALEKS math placement test scores, and scores related to the highest level of math taken in high school. The dependent variables were related to success in the first-year math and science courses, as measured by whether or not students passed those courses with a grade of C or higher. The year the student entered college, sex, and race/ethnicity were utilized as control variables.

Prior to the hierarchical logistic regression analysis that was conducted to answer the study's research questions, a hierarchical linear regression was also conducted to analyze the measure of collinearity that existed among the four independent variables (i.e., identifying those independent variables that are highly correlated with one another). Determining the level of collinearity was accomplished by analyzing the tolerance statistics, which always range from a value of 0-1 (Mertler & Vannatta, 2005). While there is no exact (agreed upon) tolerance value that determines the acceptable level of multicollinearity, typically a tolerance value of less than 0.1 serves as the statistical cutoff for significance. In other words, utilizing the results of a linear regression analysis would reveal that multicollinearity was a distinct problem for any independent variable with a tolerance value lower than 0.1 (Norusis, 1998).

My research study did not require the use of any original design instruments in order to conduct my data analysis. Instead, the study relied upon existing instruments to obtain traditional cognitive measures of students' academic preparation. Those instruments included the SAT test, ALEKS math placement test, and the use of academic coursework, grades, and overall high school GPA. With regard to the estimates of validity and reliability of these instruments, I relied on the previous studies that have been found in the literature (as described more thoroughly in Chapter II). For instance, the standardized college entrance exams such as the SAT and ACT have been utilized for decades as a measure for predicting success in college because they have been found to correlate with first year college grades at the .60 level or higher (Burton & Ramist, 2001; Camara & Echternacht, 2000). As mentioned in Chapter II, Fuertes and Sedlacek (1994) have noted that the SAT was designed "to assess the scholastic ability of high school students entering college" (p. 350). They also noted that there has been an ongoing controversy over the validity and reliability of using the SAT to predict college success. Kessel and Linn (1996) supported that notion by citing several studies that argued against using only SAT scores to predict who will and will not be successful in college. Along those lines, even though SAT scores have been traditionally utilized to predict academic success, my study aimed to compensate for any concerns over utilizing the SAT by combining those scores with other cognitive variables to predict student success.

There are several examples from the literature that support the validity of using a combination of high school performance and SAT scores to measure future success (Burton & Ramist, 2001; Camara & Echternacht, 2000; Hoffman & Lowitzki, 2005; Mathiasen, 1984; Zwick & Sklar, 2005). According to those studies and their authors' extensive review of the pertinent literature, those two variables (high school performance and SAT scores) are consistently the best predictors of success in college. Along those lines, there have been multiple studies which have shown the correlation coefficient between the (combined) variable of high school GPA/SAT score and first year college GPA to be as high as .65 to .70 (Camara & Echternacht, 2000; Patterson et al., 2012). Prior achievement in math courses has also been shown to be highly correlated with success in college-level math courses (Cox, 2000; Wheat et al., 1991). Even placement test results, which have been criticized and found to have weak predictive validity, have been found to provide more valid predictive quality when used in conjunction with other cognitive measures such as GPA (Armstrong, 2000). Based on this aspect of the literature review and the results of past studies, I am confident that the instruments that were used to measure students' academic preparation are both valid and reliable.

The strength of the research design utilized for my study lies in the fact that it utilized a regression analysis which is a subset of the type of canonical correlation that should be used for research studies that aim to find a relationship between sets of variables (Newman, 1989). My research study fell into that category because it was conducted in an effort to determine if a relationship existed between the traditional cognitive measures of an incoming college student's academic preparation and his/her ability to succeed in the math and science courses that are required by STEM degrees and their associated curricula. One weakness of my study was the fact that it only utilized cognitive variables to predict the students' academic success. While I have previously noted that those cognitive factors have been shown to be the most reliable, there are also a number of researchers who have claimed that cognitive predictors should be used in conjunction with psychological and psychosocial factors to improve the validity of predicting student success (Burton & Ramist, 2001; Kanoy et al., 1989; Wheat et al., 1991). Prior to conducting the study, there was also a concern about the possibility for the existence of multicollinearity among the four independent variables. Even when the independent variables are shown to be statistically significant, the existence of multicollinearity can be problematic for determining the amount of unique variance that each variable is contributing to the prediction of academic success in the first year of college. In order to ease those concerns, the measure of collinearity among the predictor variables was examined utilizing tolerance statistics. The results of that measure of collinearity are presented in Chapter IV.

Research and Statistical Hypotheses

RH1: The traditional cognitive measures of academic preparation (math SAT score, unweighted high school GPA, ALEKS math placement test score, and highest level of math taken in high school) significantly predict success in the first math courses taken by first-year STEM majors when controlling for year of entry (time), sex and race/ethnicity.

Full Model: PassMath = a_1U + a_2MSAT + a_3UWGPA + a_4ALEKS + a_5HighestMath + a_6Time + a_7Sex + a_8RaceEthnicity + E_1

Restricted Model: PassMath = a_1U + a_9Time + a_{10}Sex + a_{11}RaceEthnicity + E_2

RH2: Math SAT scores significantly predict success in the first math courses taken by first-year STEM majors when controlling for unweighted high school GPA, ALEKS math placement test scores, highest level of math taken in high school, year of entry (time), sex, and race/ethnicity.

Full Model: $PassMath = a_1U + a_2MSAT + a_3UWGPA + a_4ALEKS + a_5HighestMath + a_6Time + a_7Sex + a_8RaceEthnicity + E_1$

Restricted Model: $PassMath = a_1U + a_9UWGPA + a_{10}ALEKS + a_{11}HighestMath + a_{12}Time + a_{13}Sex + a_{14}RaceEthnicity + E_2$

RH3: Unweighted high school GPA significantly predicts success in the first math courses taken by first-year STEM majors when controlling for math SAT scores, ALEKS math placement test scores, highest level of math taken in high school, year of entry (time), sex, and race/ethnicity.

Full Model: $PassMath = a_1U + a_2MSAT + a_3UWGPA + a_4ALEKS + a_5HighestMath + a_6Time + a_7Sex + a_8RaceEthnicity + E_1$

Restricted Model: $PassMath = a_1U + a_9MSAT + a_{10}ALEKS + a_{11}HighestMath + a_{12}Time + a_{13}Sex + a_{14}RaceEthnicity + E_2$

RH4: ALEKS math placement test scores significantly predict success in the first math courses taken by first-year STEM majors when controlling for math SAT scores, unweighted high school GPA, highest level of math taken in high school, year of entry (time), sex, and race/ethnicity.

Full Model: $PassMath = a_1U + a_2MSAT + a_3UWGPA + a_4ALEKS + a_5HighestMath + a_6Time + a_7Sex + a_8RaceEthnicity + E_1$

Restricted Model: $PassMath = a_1U + a_9MSAT + a_{10}UWGPA + a_{11}HighestMath + a_{12}Time + a_{13}Sex + a_{14}RaceEthnicity + E_2$

RH5: The highest level of math taken in high school significantly predicts success in the first math courses taken by first-year STEM majors when controlling for math SAT scores, unweighted high school GPA, ALEKS math placement test scores, year of entry (time), sex, and race/ethnicity.

Full Model: $PassMath = a_1U + a_2MSAT + a_3UWGPA + a_4ALEKS + a_5HighestMath + a_6Time + a_7Sex + a_8RaceEthnicity + E_1$

Restricted Model: $PassMath = a_1U + a_9MSAT + a_{10}UWGPA + a_{11}ALEKS + a_{12}Time + a_{13}Sex + a_{14}RaceEthnicity + E_2$

RH6: The traditional cognitive measures of academic preparation (math SAT score, unweighted high school GPA, ALEKS math placement test score, and highest level of math taken in high school) significantly predict success in the first science courses taken by first-year STEM majors when controlling for year of entry (time), sex, and race/ethnicity.

Full Model: $PassScience = a_1U + a_2MSAT + a_3UWGPA + a_4ALEKS + a_5HighestMath + a_6Time + a_7Sex + a_8RaceEthnicity + E_1$

Restricted Model: $PassScience = a_1U + a_9Time + a_{10}Sex + a_{11}RaceEthnicity + E_2$

RH7: Math SAT scores significantly predict success in the first science courses taken by first-year STEM majors when controlling for unweighted high school GPA, ALEKS math placement test scores, highest level of math taken in high school, year of entry (time), sex, and race/ethnicity.

Full Model: $PassScience = a_1U + a_2MSAT + a_3UWGPA + a_4ALEKS + a_5HighestMath + a_6Time + a_7Sex + a_8RaceEthnicity + E_1$

Restricted Model: $PassScience = a_1U + a_9UWGPA + a_{10}ALEKS + a_{11}HighestMath + a_{12}Time + a_{13}Sex + a_{14}RaceEthnicity + E_2$

RH8: Unweighted high school GPA significantly predicts success in the first science courses taken by first-year STEM majors when controlling for math SAT scores, ALEKS math placement test scores, highest level of math taken in high school, year of entry (time), sex, and race/ethnicity.

Full Model: $PassScience = a_1U + a_2MSAT + a_3UWGPA + a_4ALEKS + a_5HighestMath + a_6Time + a_7Sex + a_8RaceEthnicity + E_1$

Restricted Model: $PassScience = a_1U + a_9MSAT + a_{10}ALEKS + a_{11}HighestMath + a_{12}Time + a_{13}Sex + a_{14}RaceEthnicity + E_2$

RH9: ALEKS math placement test scores significantly predict success in the first science courses taken by first-year STEM majors when controlling for math SAT scores, unweighted high school GPA, highest level of math taken in high school, year of entry (time), sex, and race/ethnicity.

Full Model: $PassScience = a_1U + a_2MSAT + a_3UWGPA + a_4ALEKS + a_5HighestMath + a_6Time + a_7Sex + a_8RaceEthnicity + E_1$

Restricted Model: $PassScience = a_1U + a_9MSAT + a_{10}UWGPA + a_{11}HighestMath + a_{12}Time + a_{13}Sex + a_{14}RaceEthnicity + E_2$

RH10: The highest level of math taken in high school significantly predicts success in the first science courses taken by first-year STEM majors when controlling for math SAT scores, unweighted high school GPA, ALEKS math placement test scores, year of entry (time), sex, and race/ethnicity.

Full Model: $PassScience = a_1U + a_2MSAT + a_3UWGPA + a_4ALEKS + a_5HighestMath + a_6Time + a_7Sex + a_8RaceEthnicity + E_1$

Restricted Model: $PassScience = a_1U + a_9MSAT + a_{10}UWGPA + a_{11}ALEKS + a_{12}Time + a_{13}Sex + a_{14}RaceEthnicity + E_2$

Summary

My research study utilized the institutional records of first-year STEM students who entered a large, public, urban, Hispanic-serving, research university in the Southeastern U.S. (between 2010 and 2012) to examine the relationship between certain traditional cognitive measures of their academic preparation (math SAT scores, unweighted high school GPA, ALEKS math placement test scores, and highest level of math taken in high school) and success (measured by whether or not they earned a grade of C or higher) in the math and science courses taken during their first year. Each of the students included in the study had declared that they intended to pursue one of the university's 15 STEM majors. While the institution utilized for my study has a student population that is approximately 55% female and 45% male, the sample that was utilized for my study was almost 80% male. This was attributed to the fact that the university has significantly more men enrolled in its STEM majors than women. The racial/ethnic breakdown of the sample was very comparable to the institution's demographic data. As such, almost 67% of the students were Hispanic, and the remainder was 11% White/Non-Hispanic, 11% Black/African-American, and 11% "Other."

Utilizing the information obtained from the university's student academic information system, data was collected, coded, and entered into SPSS in order to conduct a series of quantitative analyses. As those data were collected, individuals who did not have SAT scores, those who did not have ALEKS scores on file, and those who had not completed a math and science course during their first year were eliminated from the sample. My study utilized an ex-post facto research design to run a series of hierarchical logistic regression analyses on the 1018 students who were included in the sample population. Utilizing the -2*Log Likelihood statistic, full Chi-Square model, pseudo R^2 (Cox & Snell R^2), and classification model, I was able to analyze the four independent variables' statistical significance in predicting whether the students passed their first math and science courses. In doing so, my study tested each of the 10 research hypotheses in an effort to answer the research questions that were identified in Chapter I. Prior to running the logistic regression analyses a linear regression analysis was also conducted to measure the collinearity among the four independent variables. The results of the study, including the regression coefficients, statistical significance of each overall model, and the statistical significance of each individual predictor variable are provided in Chapter IV.

CHAPTER IV
RESULTS

As discussed in Chapter I, my study examined the relationship between four traditional measures of college students' academic preparation and the likelihood of passing college-level math and science courses. More specifically, the math SAT scores, unweighted high school GPAs, ALEKS (math) placement test scores, and the highest level of math taken in high school for 1018 first-year STEM majors were analyzed in relation to whether those students passed the first math and science courses they took in college. Those students all entered the same large, public, urban, Hispanic-serving, research university in the Southeastern part of the U.S. in 2010, 2011, or 2012.

The results of my study are provided below in an effort to present higher education educators with data that might be used to help predict STEM majors' likelihood for success in their first-year math and science courses. By utilizing variables such as GPA, test scores, and previous coursework to ascertain that certain STEM students are likely to fail gateway math and science courses, institutions can work toward diminishing the number of students who experience academic difficulty during their first year. Given the increased emphasis on encouraging college students to pursue STEM disciplines, my study did not aim to suggest anything about students' suitability or desirability to pursue STEM majors. Instead, the key purpose of my research involved providing an example of how institutions can utilize data that they already know about their first-year students to develop models for predicting whether their STEM majors are likely to succeed in the rigorous math and science requirements. In doing so, institutions can also utilize these data to identify alternatives (e.g., requiring students to take review courses or providing more robust supplement education opportunities) for those STEM majors who enter college with levels of academic preparation that suggest they are less likely to succeed.

Sample Population Demographics

As discussed in Chapter III, the computerized student information system of the institution utilized for my study was accessed to obtain a sample of 1367 college students who entered the institution between 2010 and 2012 and who indicated that they planned to major in one of the institution's 15 STEM majors. After removing duplicates, the initial list of students was reduced to 1018 by eliminating those who did not have SAT scores on record, those who did not have an ALEKS placement test score, and those who had not completed a math and science course during their first year. A demographic breakdown, including the year of entry, sex, and race/ethnicity, of the students has been provided below in Table 3. Even though three years of first-year students were utilized for my study, over 95% of the final sample entered the institution in 2011 or 2012. This was due in large part to the fact that the results of the ALEKS math placement test were not consistently entered into the student information system until 2011. Table 3 also highlights that almost 80% of the sample population were men (which is attributed to the fact that more men than women have traditionally pursued STEM majors) and about 67% were Hispanic (once again, the institution utilized for my study is a Hispanic-serving institution).

TABLE 3

Sample Population Demographics

	Frequency	Percentage
Entered in 2010	45	4.4%
Entered in 2011	459	45.1%
Entered in 2012	514	50.5%
Female	213	20.9%
Male	805	79.1%
Hispanic	681	66.9%
White/Non-Hispanic	113	11.1%
Black/African-American	112	11%
Other	112	11%

In addition to the demographics related to year of entry, sex, and race/ethnicity, Table 4 has been provided to illustrate the frequency of each of the 15 STEM majors within the sample population. With regard to the four areas of STEM (science, technology, engineering, and mathematics), the majority of the students (about 53%) in the sample were majoring in one of the institution's six engineering majors. The next largest group consisted of students who had selected one of the five science majors (23.2%), followed by those in the two technology majors (20.6%), and finally those who had selected one of the two mathematics degree programs (3.2%).

TABLE 4

Sample Population by STEM Major

Major	Frequency	Percentage
Science Majors	**335**	**23.2%**
Biology	130	12.8%
Chemistry	58	5.7%
Earth Science	3	0.3%
Geoscience	17	1.7%
Physics	27	2.7%
Technology Majors	**209**	**20.6%**
Computer Science	138	13.6%
Information Technology	71	7.0%
Engineering Majors	**542**	**53.2%**
Biomedical Engineering	114	11.2%
Civil Engineering	95	9.3%
Computer Engineering	94	9.2%
Electrical Engineering	52	5.1%
Environmental Engineering	23	2.3%
Mechanical Engineering	164	16.1%
Mathematics Majors	**32**	**3.2%**
Mathematics	21	2.1%
Statistics	11	1.1%

Findings

In order to conduct the statistical tests for my study, the raw data values for the four independent variables (math SAT score, unweighted high school GPA, ALEKS math placement test score, and an ordinal variable related to the highest level of math taken in high school), two dichotomous dependent variables (passed/did not pass first math course and passed/did not pass first science course), and three control variables (year the student entered college, sex, and race/ethnicity) were entered in SPSS. A binary coding method was utilized to differentiate students' year of entry, sex, and race/ethnicity. Two hierarchical linear regression analyses (one for each dependent variable) were run to measure the collinearity among the four independent variables. The results of that analysis produced the tolerance statistics that is provided below in Table 5. Although the tolerance statistic values that suggest significant multicollinearity are left up to the interpretation of the researcher, most experts agree that values of less than 0.1 indicate that multicollinearity is a distinct problem (Norusis, 1998). As noted in the table, the tolerance statistic values ranged from 0.939 to 0.994 for each variable in both of the prediction models (the one that attempted to predict those who passed math and the one that attempted to predict those who passed science). The fact that these values were much higher than the 0.1 cutoff (and, in fact, much closer to 1 than to 0) suggested that the four predictor variables were not correlated enough with one another to cause concern over utilizing all four of them in the logistic regression models.

TABLE 5

Tolerance Statistics for Both Regression Models

Predictor Variable	Passed Math?	Passed Science?
Unweighted HS GPA	0.939	0.940
Math SAT score	0.993	0.991
ALEKS score	0.994	0.994
Level of HS Math	0.992	0.991

Once the levels of collinearity were measured, the hierarchical logistic regression was conducted to determine if the independent variables were significant predictors of successfully passing math and science courses, while controlling for students' year of entry, sex, and race/ethnicity. As discussed in Chapter III, a logistic regression was selected as the appropriate statistical method because the dependent variables for my study were recorded as dichotomous nominal data (1 = passed math/science, 0 = did not pass). The remainder of this section contains each of the 10 research questions and the results of the corresponding statistical analyses.

Q1 - Do math SAT score, unweighted high school GPA, ALEKS score, and highest level of math taken in high school predict whether STEM majors will pass their first math class over and above year of entry (2010, 2011 or 2012), sex, and race/ethnicity?

Regression results indicated that the overall model of predictors was statistically significant in distinguishing between passing and failing the first math course (-2*Log Likelihood = 1210.015; χ^2 (9, N = 1018) = 196.687, $p < .001$). The -2*Log Likelihood statistic decreased from a value of 1387.828 in block one to 1210.015 in block two, which indicated that the four predictor variables accounted for a significant amount of the unexplained variance in the overall regression model. After entering the four predictor variables into the regression model, the model also went from correctly classifying 54.0% of cases on block one to 69.1% of cases on block two. In other words, by using the four predictor variables we can correctly predict whether a student passed his/her first math course 69.1% of the time as opposed to only 54% percent of the time when those four variables were not in the model. The proportion of variance accounted for by the regression model (Cox & Snell R^2) went from 1.8% (without the four predictor variables) to 17.6% (when including the predictor variables). *Wald* statistics indicated that all four of the predictor variables were significant ($p < .05$) with unweighted high school GPA (UWGPA) and highest level of high school math (HSMath) having the largest odds ratios. The regression coefficients for the four predictor variables, including UWGPA, HSMath, math SAT scores (MSAT), and math placement test scores (ALEKS), are provided below in Table 6. It is worth noting that while the ALEKS score has a statistically significant (small) effect, the negative beta weight (B-value) suggests that higher ALEKS scores do not improve the overall model's ability to predict success in math. A possible explanation for that is addressed in Chapter V. The regression coefficients for the three control variables (year of entry, sex, and race/ethnicity) were not included because none of them were statistically significant at the $p < .05$ level.

TABLE 6

Regression Coefficients for Question One

	B	*Wald*	df	*p*	*Odds Ratio*
MSAT	0.004	9.281	1	.002	1.004
UWGPA	1.880	80.104	1	.000	6.551
ALEKS	-0.010	7.953	1	.005	0.990
HSMath	0.552	22.795	1	.000	1.737
Constant	-9.130	90.773	1	.000	0.000

Note. The independent variables for this question included: math SAT score, unweighted high school GPA, ALEKS placement test score, and highest level of high school math. The dependent variable was a dichotomous variable related to whether students had passed their first math course. The control variables were year of entry, sex, and race/ethnicity.

Q2 - Does math SAT score predict whether STEM majors will pass their first math class over and above unweighted high school GPA, ALEKS score, highest level of math taken in high school, year of entry, sex, and race/ethnicity?

Regression results indicated that the overall model of predictors was statistically significant in distinguishing between passing and failing the first math course (-2*Log Likelihood = 1210.015; χ^2 (9, N = 1018) = 196.687, $p < .001$). The -2*Log Likelihood statistic decreased from a value of 1219.441 in block one to 1210.015 in block two, which indicated that math SAT score accounted for a relatively small amount of the unexplained variance in the overall regression model. The model also went from correctly classifying 69.2% of cases on block one to 69.1% of cases on block two. In other words, the ability to correctly predict whether the students passed their first math course was essentially the same with or without including the math SAT score (and was actually slightly better without using the math SAT variable). The proportion of variance accounted for by the model (Cox & Snell R^2) went from 16.8% (without math SAT score) to 17.6% (when including math SAT score). As noted previously, the *Wald* statistics indicated that the predictor variables were significant ($p < .05$) with unweighted high school GPA and highest level of high school math having the largest odds ratios. The regression coefficients for the model's predictor variables are provided below in Table 7. The regression coefficients for year of entry, sex, and race/ethnicity were not included because none of them were statistically significant at the $p < .05$ level.

TABLE 7

Regression Coefficients for Question Two

	B	*Wald*	df	*p*	*Odds Ratio*
UWGPA	1.880	80.104	1	.000	6.551
ALEKS	-0.010	7.953	1	.005	0.990
HSMath	0.552	22.795	1	.000	1.737
MSAT	0.004	9.281	1	.002	1.004
Constant	-9.130	90.773	1	.000	0.000

Note. The independent variable for this question was math SAT score. The dependent variable was a dichotomous variable related to whether students had passed their first math course. The control variables were unweighted high school GPA, ALEKS placement test score, highest level of high school math, year of entry, sex, and race/ethnicity.

Q3 - Does unweighted high school GPA predict whether STEM majors will pass their first math class over and above math SAT score, ALEKS score, highest level of math taken in high school, year of entry, sex, and race/ethnicity?

Regression results indicated that the overall model of predictors was statistically significant in distinguishing between passing and failing the first math course (-2*Log Likelihood = 1210.015; χ^2 (9, N = 1018) = 196.687, $p < .001$). The -2*Log Likelihood statistic decreased from a value of 1302.646 in block one to 1210.015 in block two, which indicated that unweighted high school GPA accounted for a significant amount of the unexplained variance in the overall regression model. The model also went from correctly classifying 62.7% of cases on block one to 69.1% of cases on block two. In other words, utilizing all of the predictor and control variables except for unweighted high school GPA provided a correct prediction of passing math 62.7% of the time, while including the unweighted high school GPA improved that prediction rate to 69.1%. The proportion of variance accounted for by the model (Cox & Snell R^2) went from 9.7% (without unweighted high school GPA) to 17.6% (when including unweighted high school GPA). As noted previously, the *Wald* statistics indicated that the predictor variables were significant ($p < .05$) with unweighted high school GPA and highest level of high school math having the largest odds ratios. The regression coefficients for the model's predictor variables are provided below in Table 8. The regression coefficients for year of entry, sex, and race/ethnicity were not included because none of them were statistically significant at the $p < .05$ level.

TABLE 8

Regression Coefficients for Question Three

	B	*Wald*	df	*p*	*Odds Ratio*
MSAT	0.004	9.281	1	.002	1.004
ALEKS	-0.010	7.953	1	.005	0.990
HSMath	0.522	22.795	1	.000	1.737
UWGPA	1.880	80.104	1	.000	6.551
Constant	-9.130	90.773	1	.000	0.000

Note. The independent variable for this question was unweighted high school GPA. The dependent variable was a dichotomous variable related to whether students had passed their first math course. The control variables were math SAT score, ALEKS placement test score, highest level of high school math, year of entry, sex, and race/ethnicity.

Q4 - Does ALEKS score predict whether STEM majors will pass their first math class over and above math SAT score, unweighted high school GPA, highest level of math taken in high school, year of entry, sex, and race/ethnicity?

Regression results indicated that the overall model of predictors was statistically significant in distinguishing between passing and failing the first math course (-2*Log Likelihood = 1210.015; χ^2 (9, N = 1018) = 196.687, $p < .001$). The -2*Log Likelihood statistic decreased from a value of 1218.108 in block one to 1210.015 in block two, which indicated that ALEKS score accounted for a relatively small amount of the unexplained variance in the overall regression model. The model also correctly classified 69.1% of cases on both block one and block two. In other words, the ability to correctly predict whether the students passed their first math course was statistically the same with or without including the ALEKS score. The proportion of variance accounted for by the model (Cox & Snell R^2) went from 16.9% (without ALEKS score) to 17.6% (when including ALEKS score). As noted previously, the *Wald* statistics indicated that the predictor variables were significant ($p < .05$) with unweighted high school GPA and highest level of high school math having the largest odds ratios. The regression coefficients for the model's predictor variables are provided below in Table 9. The regression coefficients for year of entry, sex, and race/ethnicity were not included because none of them were statistically significant at the $p < .05$ level.

TABLE 9

Regression Coefficients for Question Four

	B	*Wald*	df	*p*	*Odds Ratio*
MSAT	0.004	9.281	1	.002	1.004
UWGPA	1.880	80.104	1	.000	6.551
HSMath	0.552	22.795	1	.000	1.737
ALEKS	-0.010	7.953	1	.005	0.990
Constant	-9.130	90.773	1	.000	0.000

Note. The independent variable for this question was ALEKS placement test score. The dependent variable was a dichotomous variable related to whether students had passed their first math course. The control variables were math SAT score, unweighted high school GPA, highest level of high school math, year of entry, sex, and race/ethnicity.

Q5 - Does highest math taken in high school predict whether STEM majors will pass their first math class over and above math SAT score, unweighted high school GPA, ALEKS score, year of entry, sex, and race/ethnicity?

Regression results indicated that the overall model of predictors was statistically significant in distinguishing between passing and failing the first math course (-2*Log Likelihood = 1210.015; χ^2 (9, N = 1018) = 196.687, $p < .001$). The -2*Log Likelihood statistic decreased from a value of 1233.696 in block one to 1210.015 in block two, which indicated that highest math taken in high school accounted for a moderate amount of the unexplained variance in the overall regression model. The model also went from correctly classifying 67.7% of cases on block one to 69.1% of cases on block two. In other words, utilizing all of the predictor and control variables except for highest math taken in high school provided a correct prediction of passing math 67.7% of the time, while including the highest math taken in high school slightly improved that prediction rate to 69.1%.The proportion of variance accounted for by the model (Cox & Snell R^2) went from 15.6% (without highest level of math taken in high school) to 17.6% (when including highest level of math taken in high school). As noted previously, the *Wald* statistics indicated that the predictor variables were significant ($p < .05$) with unweighted high school GPA and highest level of high school math having the largest odds ratios. The regression coefficients for the model's predictor variables are provided below in Table 10. The regression coefficients for year of entry, sex, and race/ethnicity were not included because none of them were statistically significant at the $p < .05$ level.

TABLE 10

Regression Coefficients for Question Five

	B	*Wald*	df	*p*	*Odds Ratio*
MSAT	0.004	9.281	1	.002	1.004
UWGPA	1.880	80.104	1	.000	6.551
ALEKS	-0.010	7.953	1	.005	0.990

HSMath	0.552	22.795	1	.000	1.737
Constant	-9.130	90.773	1	.000	0.000

Note. The independent variable for this question was highest level of high school math. The dependent variable was a dichotomous variable related to whether students had passed their first math course. The control variables were math SAT score, unweighted high school GPA, ALEKS placement test score, year of entry, sex, and race/ethnicity.

Q6 - Do math SAT score, unweighted high school GPA, ALEKS score, and highest level of math taken in high school predict whether STEM majors will pass their first science class over and above year of entry (2010, 2011 or 2012), sex, and race/ethnicity?

Regression results indicated that the overall model of predictors was statistically significant in distinguishing between passing and failing the first science course (-2*Log Likelihood = 1027.470; χ^2 (9, N = 1018) = 176.924, $p < .001$). The -2*Log Likelihood statistic decreased from a value of 1198.983 in block one to 1027.470 in block two, which indicated that the four predictor variables accounted for a significant amount of the unexplained variance in the overall regression model. The model also went from correctly classifying 66.6% of cases on block one to 70.7% of cases on block two. In other words, by using the four predictor variables we can correctly predict whether a student passed his/her first science course 70.7% of the time (as opposed to only 66.6% percent of the time without using those four variables). The proportion of variance accounted for by the model (Cox & Snell R^2) went from 0.6% (without the four predictor variables) to 17.1% (when including the predictor variables). *Wald* statistics indicated that the only predictor variable that was significant ($p < .05$) for predicting that students would pass their first science course was unweighted high school GPA. The regression coefficients for each predictor variable are provided below in Table 11. The regression coefficients for the three control variables (year of entry, sex, and race/ethnicity) were not included because none of them were statistically significant at the $p < .05$ level.

TABLE 11

Regression Coefficients for Question Six

	B	*Wald*	df	*p*	*Odds Ratio*
MSAT	0.002	2.764	1	.096	1.002
UWGPA	2.366	96.333	1	.000	10.650
ALEKS	0.005	2.082	1	.149	1.005
HSMath	0.128	1.043	1	.307	1.136
Constant	-9.076	72.418	1	.000	0.000

Note. The independent variables for this question included: math SAT score, unweighted high school GPA, ALEKS placement test score, and highest level of high school math. The dependent variable was a dichotomous variable related to whether students had passed their first science course. The control variables were year of entry, sex, and race/ethnicity.

Q7 - Does math SAT score predict whether STEM majors will pass their first science class over and above unweighted high school GPA, ALEKS score, highest level of math taken in high school, year of entry, sex, and race/ethnicity?

Regression results indicated that the overall model of predictors was statistically significant in distinguishing between passing and failing the first science course (-2*Log Likelihood = 1027.470; χ^2 (9, N = 1018) = 176.924, $p < .001$). The -2*Log Likelihood statistic decreased from a value of 1030.244 in block one to 1027.47 in block two, which indicated that math SAT score accounted for a relatively small amount of the unexplained variance in the overall regression model. The model also went from correctly classifying 70.9% of cases on block one to 70.7% of cases on block two. In other words, the ability to correctly predict whether the students passed their first science course was essentially the same with or without including the math SAT score (and was actually slightly better without using the math SAT variable). The proportion of variance accounted for by the model (Cox & Snell R^2) went from 16.8% (without math SAT score) to 17.1% (when including math SAT score). As noted previously, the *Wald* statistics indicated that the only predictor variable that was significant ($p < .05$) for predicting that students would pass their first science course was unweighted high school GPA. The regression coefficients for the model's predictor variables are provided below in Table 12. The regression coefficients for year of entry, sex, and race/ethnicity were not included because none of them were statistically significant at the $p < .05$ level.

TABLE 12

Regression Coefficients for Question Seven

	B	*Wald*	df	*p*	*Odds Ratio*
UWGPA	2.366	96.333	1	.000	10.650
ALEKS	0.005	2.082	1	.149	1.005
HSMath	0.128	1.043	1	.307	1.136
MSAT	0.002	2.764	1	.096	1.002
Constant	-9.076	72.418	1	.000	0.000

Note. The independent variable for this question was math SAT score. The dependent variable was a dichotomous variable related to whether students had passed their first science course. The control variables were unweighted high school GPA, ALEKS placement test score, highest level of high school math, year of entry, sex, and race/ethnicity.

Q8 - Does unweighted high school GPA predict whether STEM majors will pass their first science class over and above math SAT score, ALEKS score, highest level of math taken in high school, year of entry, sex, and race/ethnicity?

Regression results indicated that the overall model of predictors was statistically significant in distinguishing between passing and failing the first science course (-2*Log Likelihood = 1027.470; χ^2 (9, N = 1018) = 176.924, $p < .001$). The -2*Log Likelihood statistic decreased from a value of 1144.600 in block one to 1027.470 in block two, which indicated that unweighted high school GPA accounted for a significant amount of the unexplained variance in the overall regression model. The model also went from correctly classifying 66.7% of cases on block one to 70.7% of cases on block two. In other words, utilizing all of the predictor and control variables except for unweighted high school GPA provided a correct prediction of passing science 66.7% of the time, while including the unweighted high school GPA improved that prediction rate to 70.7%. The proportion of variance accounted for by the model (Cox & Snell R^2) went from 6.1% (without unweighted high school GPA) to 17.1% (when including unweighted high school GPA). As noted previously, the *Wald* statistics indicated that the only predictor variable that was significant ($p < .05$) for predicting that students would pass their first science course was unweighted high school GPA. The regression coefficients for the model's predictor variables are provided below in Table 13. The regression coefficients for year of entry, sex, and race/ethnicity were not included because none of them were statistically significant at the $p < .05$ level.

TABLE 13

Regression Coefficients for Question Eight

	B	*Wald*	df	*p*	*Odds Ratio*
MSAT	0.002	2.764	1	.096	1.002
ALEKS	0.005	2.082	1	.149	1.005
HSMath	0.128	1.043	1	.307	1.136
UWGPA	2.366	96.333	1	.000	10.650
Constant	-9.076	72.418	1	.000	0.000

Note. The independent variable for this question was unweighted high school GPA. The dependent variable was a dichotomous variable related to whether students had passed their first science course. The control variables were math SAT score, ALEKS placement test score, highest level of high school math, year of entry, sex, and race/ethnicity.

Q9 - Does ALEKS score predict whether STEM majors will pass their first science class over and above math SAT score, unweighted high school GPA, highest level of math taken in high school, year of entry, sex, and race/ethnicity?

Regression results indicated that the overall model of predictors was statistically significant in distinguishing between passing and failing the first science course (-2*Log Likelihood = 1027.470; χ^2 (9, N = 1018) = 176.924, $p < .001$). The -2*Log Likelihood statistic decreased from a value of 1029.556 in block one to 1027.470 in block two, which indicated that ALEKS score accounted for a relatively small amount of the unexplained variance in the overall regression model. The model also went from correctly classifying 71.5% of cases on block one to 70.7% of cases on block two. In other words, utilizing all of the predictor and control variables except for ALEKS score provided a correct prediction of passing science 71.5% of the time, while including the ALEKS score actually reduced that prediction rate to 70.7%. The proportion of variance accounted for by the model (Cox & Snell R^2) went from 16.9% (without ALEKS score) to 17.1% (when including ALEKS score). As noted previously, the *Wald* statistics indicated that the only predictor variable that was significant ($p < .05$) for predicting that students would pass their first science course was unweighted high school GPA. The regression coefficients for the model's predictor variables are provided below in Table 14. The regression coefficients for year of entry, sex, and race/ethnicity were not included because none of them were statistically significant at the $p < .05$ level.

TABLE 14

Regression Coefficients for Question Nine

	B	*Wald*	df	*p*	*Odds Ratio*
UWGPA	2.366	96.333	1	.000	10.650
MSAT	0.002	2.764	1	.096	1.002
HSMath	0.128	1.043	1	.307	1.136
ALEKS	0.005	2.082	1	.149	1.005
Constant	-9.076	72.418	1	.000	0.000

Note. The independent variable for this question was ALEKS placement test score. The dependent variable was a dichotomous variable related to whether students had passed their first science course. The control variables were math SAT score, unweighted high school GPA, highest level of high school math, year of entry, sex, and race/ethnicity.

Q10 - Does highest math taken in high school predict whether STEM majors will pass their first science class over and above math SAT score, unweighted high school GPA, ALEKS score, year of entry, sex, and race/ethnicity?

Regression results indicated that the overall model of predictors was statistically significant in distinguishing between passing and failing the first science course (-2*Log Likelihood = 1027.470; χ^2 (9, N = 1018) = 176.924, p < .001). The -2*Log Likelihood statistic decreased from a value of 1028.515 in block one to 1027.470 in block two, which indicated that highest math taken in high school accounted for a relatively small amount of the unexplained variance in the overall regression model. The model also went from correctly classifying 70.5% of cases on block one to 70.7% of cases on block two. In other words, the ability to correctly predict whether the students passed their first science course was only slightly better when including the highest level of math taken in high school (but was essentially the same with or without that variable). The proportion of variance accounted for by the model (Cox & Snell R^2) went from 17.0% (without highest level of math taken in high school) to 17.1% (when including highest level of math taken in high school). As noted previously, the *Wald* statistics indicated that the only predictor variable that was significant (p < .05) for predicting that students would pass their first science course was unweighted high school GPA. The regression coefficients for the model's predictor variables are provided below in Table 15. The regression coefficients for year of entry, sex, and race/ethnicity were not included because none of them were statistically significant at the p < .05 level.

TABLE 15

Regression Coefficients for Question Ten

	B	*Wald*	df	*p*	*Odds Ratio*
MSAT	0.002	2.764	1	.096	1.002
UWGPA	2.366	96.333	1	.000	10.650
ALEKS	0.005	2.082	1	.149	1.005
HSMath	0.128	1.043	1	.307	1.136
Constant	-9.076	72.418	1	.000	0.000

Note. The independent variable for this question was highest level of high school math. The dependent variable was a dichotomous variable related to whether students had passed their first science course. The control variables were math SAT score, unweighted high school GPA, ALEKS placement test score, year of entry, sex, and race/ethnicity.

Summary

As previously noted, my study examined data on 1018 first-year STEM majors (who entered a large, public, Hispanic-serving, research university) in an effort to predict whether they passed their first math and science courses. Ninety-five percent of the students entered college in 2011 or 2012, 80% were men, 67% were Hispanic, and 53% were pursuing engineering majors. The four predictor variables, math SAT score, unweighted high school GPA, ALEKS math placement test score, and the highest level of math taken in high school, were utilized to develop models for predicting the students' success in math and science. Utilizing a hierarchical linear regression analysis, the resulting tolerance statistics were utilized to measure the level of collinearity among those four predictor variables. After reviewing those statistics, it was determined that there was no significant multicollinearity present in either of the overall prediction models.

This chapter also provided the results of the hierarchical logistic regression analyses that were conducted to answer my study's 10 research questions. It was noted that the overall model of predictors was statistically significant for predicting students who passed math (χ^2 (9, N = 1018) = 196.687, $p < .001$) as well as for predicting students who passed science (χ^2 (9, N = 1018) = 176.924, $p < .001$). The -2*Log Likelihood statistics also indicated that the four predictor variables accounted for a significant amount of unexplained variance in the overall regression models for predicting success in both math and science. Along those lines, the Pseudo R^2 statistics (Cox & Snell R^2) indicated that the models that included all four predictor variables accounted for a greater proportion of the existing variance. For the models that utilized "passing math" as the dependent variable, the four predictor variables accounted for 17.6% of the proportion of variance (as opposed to 1.8% without those predictors). For the models that utilized “passing science” as the dependent variable, the four predictor variables accounted for 17.1% of the proportion of variance (as opposed to 0.6% without those predictors). In both sets of models, the unweighted high school GPA variable accounted for the largest proportion of variance.

The analyses that utilized “passing math” as the dependent variable also resulted in *Wald* statistics that indicated that all four predictor variables were significant ($p < .05$) for predicting those students who passed math. The classification model for those analyses revealed that utilizing all four predictor variables improved the ability to correctly predict whether students would pass math from 54% to 69.1%. For the analyses that utilized “passing science” as the dependent variable, the *Wald* statistics indicated that unweighted high school GPA was the only significant ($p < .05$) variable. Utilizing all four variables for that classification model improved the ability to correctly predict whether students would pass science from 66.6% to 70.7%. An interpretation of all of these results and how they can be utilized to influence future research and institutional practices will be discussed in greater detail in Chapter V.

CHAPTER V
DISCUSSION

The results of my research study supported my initial hypothesis that first-year STEM majors' traditional cognitive measures of academic preparation can be used to help predict whether they are likely to pass key STEM-related courses. Those traditional cognitive measures served as the predictor (independent) variables for my study and included math SAT score, unweighted high school GPA, ALEKS math placement test score, and the highest level of math taken in high school. The two dependent variables for my study were dichotomous variables that indicated whether or not each student had passed his/her first math course (utilized as the dependent variable for the first five research questions) and whether or not each student had passed his/her first science course (utilized as the dependent variable for the remaining five research questions). The four predictor variables were analyzed, utilizing hierarchical logistic regression, first as a group and then individually to determine how significant they were for predicting those students who passed their first math course. Those analyzes were then repeated to determine how significant the predictor variables (as a group and individually) were for predicting those students who passed their first science course. In order to conduct those analyses and test the overarching hypothesis that traditional cognitive measures of academic preparation were significant for predicting success in math and science, the following 10 research questions were developed:

- Q1 - Do traditional cognitive measures of academic preparation (math SAT scores, unweighted high school GPA, math placement test scores, and highest level of math taken in high school) significantly differentiate between STEM students who pass math and those who fail math when controlling for year of entry, sex, and race/ethnicity?

- Q2 - Does a freshman STEM major’s math SAT score account for a significant amount of unique variance when predicting success in the first math course taken when controlling for unweighted high school GPA, math placement test score, highest level of math taken in high school, year of entry, sex, and race/ethnicity?

- Q3 - Does a freshman STEM major’s unweighted high school GPA account for a significant amount of unique variance when predicting success in the first math course taken when controlling for math SAT score, math placement test score, highest level of math taken in high school, year of entry, sex, and race/ethnicity?

- Q4 - Does a freshman STEM major’s math placement test score account for a significant amount of unique variance when predicting success in the first math course taken when controlling for math SAT score, unweighted high school GPA, highest level of math taken in high school, year of entry, sex, and race/ethnicity?

- Q5 - Does a freshman STEM major’s highest level of math taken in high school account for a significant amount of unique variance when predicting success in the first math course taken when controlling for math SAT score, unweighted high school GPA, math placement test score, year of entry, sex, and race/ethnicity?

- Q6 - Do traditional cognitive measures of academic preparation (math SAT scores, unweighted high school GPA, math placement test scores, and highest level of math taken in high school) significantly differentiate between STEM students who pass science and those who fail science when controlling for year of entry, sex, and race/ethnicity?

- Q7 - Does a freshman STEM major’s math SAT score account for a significant amount of unique variance when predicting success in the first science course taken when controlling for unweighted high school GPA, math placement test score, highest level of math taken in high school, year of entry, sex, and race/ethnicity?

- Q8 - Does a freshman STEM major's unweighted high school GPA account for a significant amount of unique variance when predicting success in the first science course taken when controlling for math SAT score, math placement test score, highest level of math taken in high school, year of entry, sex, and race/ethnicity?
- Q9 - Does a freshman STEM major's math placement test score account for a significant amount of unique variance when predicting success in the first science course taken when controlling for math SAT score, unweighted high school GPA, highest level of math taken in high school, year of entry, sex, and race/ethnicity?
- Q10 - Does a freshman STEM major's highest level of math taken in high school account for a significant amount of unique variance when predicting success in the first science course taken when controlling for math SAT score, unweighted high school GPA, math placement test score, year of entry, sex, and race/ethnicity?

The academic records of the first-year STEM students who had entered a large, public, urban, Hispanic-serving, research university between 2010 and 2012 were reviewed to comprise the data for my study. After eliminating those individuals who did not possess values for each of the independent and dependent variables, a sample of 1018 students was utilized to conduct the statistical analyses. The detailed findings and interpretations of those analyses are provided in the next section of this chapter. In summary (as discussed in Chapter IV), the analyses that were utilized to answer the first five research questions resulted in an overall model that was statistically significant for predicting those students who passed their first math course. With regard to predicting those who passed math, all four of the predictor variables were statistically significant. Along those lines, the regression models that measured the effect of all four variables together (in the analysis for research question one), the unweighted high school GPA variable (in the analysis for research question three), and the variable for highest math taken in high school (in the analysis for research question five) provided the most predictive power. The data that were presented in the classification model indicated that the use of those particular variables improved the ability to correctly predict if students passed math by up to as much as 15%. Conversely, the data also indicated that including the math SAT score and the ALEKS score did not result in any improvement in predicting those students who passed their first math course (and in the case of the math SAT, its inclusion slightly lowered the ability to correctly predict those who passed math).

The analyses that were utilized to answer the last five research questions indicated that the overall model was also statistically significant for predicting those students who passed their first science course. Further analysis revealed, however, that (individually) the only predictor variable that was statistically significant for predicting whether the students passed their first science course was the unweighted high school GPA variable. The classification model data indicated that the only models that provided any improvement in correctly predicting those who passed science were the one that measured the effect of all four predictor variables (in the analysis for research question six) and the one that measured the effect of unweighted high school GPA (in the analysis for research question eight). In both of those models, the ability to correctly predict those who passed science was improved by about 4%. The models that analyzed the effect of the other three predictor variables either provided virtually no improvement (in the case of highest math taken in high school) or actually lowered the ability to correctly predict those who had passed their first science course (in the cases of math SAT score and ALEKS score).

As discussed previously, my research study was conducted to complement the existing research related to predicting first-year success in college and the factors associated with succeeding in the STEM disciplines. With regard to first-year success, the literature has supported the notion that the first six months to one year of college are the most critical to retaining students and predicting their eventual graduation (Siegel, 2011; Stovall, 2000). Additionally, past research has found that both retention and student success have more to do with the attributes and academic preparation of the students themselves rather than institutional factors (Freeman et al., 2007). Placing students into appropriate courses that align with their abilities and academic preparation has also been found to be a best practice by those who are working to improve college student success (Kim et al., 2010).

My interest in utilizing key cognitive and academic attributes to ensure that first-year students take courses in which they are likely to succeed led directly to the development of my research study. My decision to focus on first-year STEM majors was motivated by the increased emphasis within the U.S. educational system on preparing more students to enter STEM disciplines (Sadler et al., 2012). As Thompson and Bolin (2011) noted, not only do U.S. institutions need to improve the graduation rates of their STEM majors but the highest dropout rates for STEM students have occurred during the first year of college. Along those lines, my research study was conducted with the hope that the results can be utilized to inform institutional policies regarding the placement of first-year STEM majors into math and science courses that align with their likelihood for success. My findings suggest that the most critical attributes for predicting the success of those first-year STEM majors are high school GPA and the courses they took in high school. As such, the results of my study can serve as a model for assisting institutions with identifying students who might benefit from additional resources and support in order to increase their chances of succeeding in the rigorous math and science courses that are required of STEM majors. My results also suggest that colleges and universities should be doing more to work with K-12 educators to inform their efforts to prepare students for the rigors associated with pursuing STEM degrees.

Findings and Interpretations

In order to test the 10 research hypotheses that were developed for my study, I utilized a set of quantitative analyses that revealed the relationship between the four cognitive (independent) variables and the ability to predict whether students passed their first math and science course. As mentioned previously, hierarchical logistic regression analyses were conducted to obtain the data that were presented in Chapter IV. Those data will once again be presented in this section along with a more detailed interpretation of what they tell us about the relationship between the independent and dependent variables.

Prior to conducting the logistic regression analyses that were utilized to test the significance of the four predictor variables, I first tested the measure of collinearity among those four variables. That was accomplished by running a hierarchical linear regression for each of the study's two dependent variables. Those regression analyses provided the tolerance statistics that were needed to determine if multicollinearity was a distinct problem. According to Norusis (1998), most researchers agree that a tolerance level of less than 0.1 indicates that two or more of the predictor variables are highly correlated. While including highly correlated predictor variables does not reduce the predictive power of the overall model, it can impact the validity of determining the impact of each individual predictor variable (McNeil et al., 2012). For my study, the tolerance statistics for each independent variable were all much higher than the 0.1 cutoff for both of the regression models (for predicting success in math and for predicting success in science). As a result, I concluded that the measure of collinearity among the four independent variables was not high enough to warrant removing any of the predictors from the model.

It was noted in Chapter IV that the overall model of predictors was statistically significant in predicting those students who passed their first math course (-2*Log Likelihood = 1210.015; χ^2 (9, N =1018) = 196.687, $p < .001$) and in predicting those students who passed their first science course (-2*Log Likelihood = 1027.470; χ^2 (9, N = 1018) = 176.924, $p < .001$). Both the -2*Log Likelihood and the χ^2 statistic provided a measure of whether the model that contained the four predictor variables had reduced the amount of the unexplained variance that had been present in the base model (i.e., the model that contained only the control variables). Since the alpha level for my study was set at $p < .05$, the resulting significance value ($p = .000$, which has been interpreted and reported as $p < .001$ throughout my results and discussion) indicated that the model was indeed statistically significant. In other words, the presence of the four predictor variables that were included in the overall model for my study explained a significant amount of the original variability in the data.

With regard to each of the four predictor variables, there were a number of statistics that were used to interpret the significance they had on the full regression models. Those statistics included the regression coefficients (B), *Wald* statistic, significance value (p), and the Odds Ratio (e^B) for each of the predictor variables. The values for each of those statistics were provided in Tables 6-15 as part of Chapter IV. For the first five research questions, which related to predicting those students who had passed math, the *Wald* statistics and the corresponding *p*-values indicated that all four of the predictor variables were statistically significant. More specifically, the unweighted high school GPA and highest math taken in high school were both significant at the $p < .001$ level, with *Wald* values of 80.104 and 22.795 respectively. The *Wald* statistics for math SAT score (9.281) and ALEKS math placement test score (7.953) were also significant at the $p = .002$ and $p = .005$ levels respectively. The values for the regression coefficients (B) and the Odds Ratio provided another measure of how each independent variable was related to the model's ability to correctly predict those who had passed math. With regard to the regression coefficients, one interesting result (that was pointed out in Chapter IV) was that the B-value for the ALEKS placement test score ($B = -.010$) was both statistically significant and also a negative number. What this indicated was that even though the effect was small (since the value was close to zero), as students' ALEKS scores increased the weight or impact on the prediction model slightly decreased in relation to the other variables. At first glance this might seem counterintuitive since those students with higher ALEKS scores would seem to have more math knowledge and ability. One possible explanation for that result is that those students with a higher ALEKS score were eligible to enroll in higher level math courses than those students who scored lower on the ALEKS assessment. In other words, a higher ALEKS score often placed students into math courses that involve greater academic rigor and are therefore more difficult to pass. Another possible explanation is that the ALEKS assessment does not do a sufficient job of placing students into the appropriate math course. With regard to unweighted high school GPA, a *B*-value of 1.880 translated to an Odds Ratio of 6.551, meaning that for every one unit increase in GPA, the model was 6.551 times more likely to correctly predict those students who had passed math. Likewise, the Odds Ratios for the other three variables indicated that for every one unit increase in each, the model was 1.737 times (for highest level of high school math), 1.004 times (for math SAT score), and 0.99 times (for ALEKS score) more likely to correctly predict those who passed math. Even though all of the predictors were statistically significant, the Odds Ratios for math SAT and ALEKS indicated that those scores did not really improve the chances of correctly predicting the outcome of passing math. Overall, the unweighted high school GPA and highest level of math taken in high school were the best predictors of those students who had passed their first math course. The predictive power of each of these variables is described and interpreted in more detail later in this section.

For the last five research questions, which related to predicting those students who had passed science, the *Wald* statistics and the corresponding *p*-values indicated that unweighted high school GPA was the only predictor variable that was significant to predicting success in science (*Wald* = 96.333, $p < .001$). For those five regression models, the unweighted high school GPA variable had a *B*-value of 2.366 and an Odds Ratio of 10.65. In other words, for every one unit increase in GPA the model was 10.65 times more likely to correctly predict those students who had passed their first science course. Since none of the other variables utilized in my study were significant at the $p < .05$ level, their regression coefficients and Odds Ratios were statistically irrelevant. Once again, the predictive power of each independent variable is described further (below) as part of the discussion on each research hypothesis.

The other statistics that were utilized to address the research hypotheses included the Pseudo R^2 (Cox & Snell R^2) statistic, which provided the proportion of variance that could be explained by adding variables to the model, and the classification model, which provided the percentages of correctly predicting those who had passed math or science that were accounted for by including the various predictor variables. The results and interpretations for these two statistics are included below as part of the discussion on whether I rejected or failed to reject each of the null research hypotheses.

Null Hypothesis 1: The traditional cognitive measures of academic preparation (math SAT score, unweighted high school GPA, ALEKS math placement test score, and highest level of math taken in high school) do not significantly predict success in the first math courses taken by first-year STEM majors when controlling for year of entry, sex, and race/ethnicity.

As the overall model of predictors for the dependent variable that related to passing math, the resulting statistics (-2*Log Likelihood = 1210.015; χ^2 (9, N =1018) = 196.687, $p < .001$) enabled me to reject this null hypothesis. Since those values were indeed significant, we can conclude that the predictor variables reduced the amount of unexplained variance that had been present in the base model. The fact that the -2*Log Likelihood statistic decreased by a value of 177.813 after including the four predictor variables also showed that those variables accounted for a significant amount of the unexplained variance in the overall regression model. Once the four predictor variables were entered into the regression model, the model also went from correctly classifying 54.0% of cases to correctly classifying 69.1% of cases. In other words, by using the four predictor variables the model correctly predicted whether a student passed his/her first math course 69.1% of the time (as opposed to only 54% percent of the time when those four variables were not included in the model). In addition, the proportion of variance accounted for by the regression model (Cox & Snell R^2) went from 1.8% (without the four predictor variables) to 17.6% (when including the predictor variables). The fact that an additional 15.8% of the variance was accounted for by the full model clearly supports the use of these variables as a means for predicting success in first-year STEM majors' first math course. This result supported the literature that has emphasized the importance of high school preparation and that has found students' pre-college academic attributes useful for predicting college grades (Cole & Espinosa, 2008; Veenstra et al., 2008; Williford, 2009).

Null Hypothesis 2: Math SAT scores do not significantly predict success in the first math courses taken by first-year STEM majors when controlling for unweighted high school GPA, ALEKS math placement test scores, highest level of math taken in high school, year of entry, sex, and race/ethnicity.

Since the statistics for the overall model of predictors indicated that this predictor variable was significant (*Wald* = 9.281, p = .002), we must once again reject this null hypothesis. Even though this variable was significant to that overall model, there were additional results that indicated that this variable did not contribute much toward the effort to predict those who had passed their first math course. For example, the -2*Log Likelihood statistic only decreased by a value of 9.426, which indicated that math SAT score accounted for a relatively small amount of the unexplained variance in the overall regression model. The model also went from correctly classifying 69.2% of cases to correctly classifying 69.1% of cases. In other words, the ability to correctly predict whether the students passed their first math course was essentially the same with or without the math SAT score (and was actually slightly better without the math SAT variable). The proportion of variance accounted for by the model (Cox & Snell R^2) went from 16.8% (without math SAT score) to 17.6% (when including math SAT score), an increase of less than

1%. All of these data suggest that despite its statistical significance this variable had very little predictive power. While this result would support those who criticize the use of the SAT, it goes against those studies which have found statistically significant correlations between SAT scores and first-year college grades/GPA (Burton & Ramist, 2001; Camara & Echternacht, 2000).

Null Hypothesis 3: Unweighted high school GPA does not significantly predict success in the first math courses taken by first-year STEM majors when controlling for math SAT scores, ALEKS math placement test scores, highest level of math taken in high school, year of entry, sex, and race/ethnicity.

Since the statistics for the overall model of predictors indicated that this predictor variable was significant (*Wald* = 80.104, $p < .001$), we must once again reject this null hypothesis. In addition to this variable's significance for the overall model, there were also additional results that indicated that this variable contributed significantly toward the effort to predict those who had passed their first math course. For example, the fact that the -2*Log Likelihood statistic decreased by a value of 92.631 after including the unweighted high school GPA variable demonstrated that it accounted for a significant amount of the unexplained variance in the overall regression model. The model also went from correctly classifying 62.7% of cases to correctly classifying 69.1% of cases. In other words, utilizing all of the predictor and control variables except for unweighted high school GPA provided a correct prediction of passing math 62.7% of the time, while including the unweighted high school GPA improved that prediction rate to 69.1%. The proportion of variance accounted for by the model (Cox & Snell R^2) went from 9.7% (without unweighted high school GPA) to 17.6% (when including unweighted high school GPA). All of these data, when compared to the results of the other analyses, indicated that this variable had more predictive power than any other individual variable. This result supported several studies from the literature that have found high school grades to be the best predictor of the grades students will earn in college (Belfield & Crosta, 2012; Chase & Jacobs, 1989;

Hoffman & Lowitzki, 2005; Nagashi & Slade, 2012; Thompson & Bolin, 2011; Williford, 2009; Zwick & Sklar, 2005).

Null Hypothesis 4: ALEKS math placement test scores do not significantly predict success in the first math courses taken by first-year STEM majors when controlling for math SAT scores, unweighted high school GPA, highest level of math taken in high school, year of entry, sex, and race/ethnicity.

Since the statistics for the overall model of predictors indicated that this predictor variable was significant (*Wald* = 7.953, p = .005), we must once again reject this null hypothesis. Even though this variable was significant to that overall model, there were additional results that indicated that this variable did not contribute much toward the effort to predict those who had passed their first math course. For example, the -2*Log Likelihood statistic only decreased by a value of 8.093, which indicated that the ALEKS score accounted for a relatively small amount of the unexplained variance in the overall regression model. The model also correctly classified 69.1% of cases on both block one and block two. In other words, the ability to correctly predict whether the students passed their first math course was statistically the same with or without the ALEKS score. The proportion of variance accounted for by the model (Cox & Snell R^2) went from 16.9% (without ALEKS score) to 17.6% (when including ALEKS score), an increase of less than 1%. As mentioned previously, the (statistically significant) negative *B*-value for this variable also indicated that as the students' ALEKS scores increased the impact of those scores were slightly diminished in relation to the other predictor variables in the model. Once again, this was possibly due to the fact that those with higher ALEKS scores attempted more rigorous math courses during their first year than those with lower ALEKS scores. All of this suggests that despite its statistical significance this variable had very little predictive power. While that raises questions about the effectiveness of the ALEKS test itself, it also goes against what I found in the literature regarding the use of placement tests. As I noted in Chapter II, there have been a number of studies that have found placement tests to be useful for predicting college students who will pass math courses as well as predicting their overall first-year GPA (Cox, 2000; Scott-Clayton, 2012; Veenstra et al., 2008; Wheat et al., 1991). Based on my results, that did not hold true for the students in my study.

Null Hypothesis 5: The highest level of math taken in high school does not significantly predict success in the first math courses taken by first-year STEM majors when controlling for math SAT scores, unweighted high school GPA, ALEKS math placement test scores, time, sex, and race/ethnicity.

Since the statistics for the overall model of predictors indicated that this predictor variable was significant (*Wald* = 22.795, $p < .001$), we must once again reject this null hypothesis. In addition to this variable's significance for the overall model, there were also additional results that indicated that this variable contributed significantly toward the effort to predict those who had passed their first math course. For example, the fact that the -2*Log Likelihood statistic decreased by a value of 23.681 after including the variable for highest level of math taken in high school demonstrated that this variable accounted for a moderate amount of the unexplained variance in the overall regression model. The model also went from correctly classifying 67.7% of cases to correctly classifying 69.1% of cases. In other words, utilizing all of the predictor and control variables except for highest math taken in high school provided a correct prediction of passing math 67.7% of the time, while including the highest math taken in high school improved that prediction rate to 69.1%. The proportion of variance accounted for by the model (Cox & Snell R^2) went from 15.6% (without highest level of math taken in high school) to 17.6% (when including highest level of math taken in high school). All of these data, when compared to the results of the other analyses, indicated that this variable had more predictive power than any other individual variable except for unweighted high school GPA. That supported studies from the literature that have found students' high school math preparation to be a significant factor for predicting success in college math courses (Hemmings et al., 2011; Kessel & Linn, 1996; Levin & Wyckoff, 1991; Sadler et al., 2012; Veenstra et al., 2008).

Null Hypothesis 6: The traditional cognitive measures of academic preparation (math SAT score, unweighted high school GPA, ALEKS math placement test score, and highest level of math taken in high school) do not significantly predict success in the first science courses taken by first-year STEM majors when controlling for year of entry, sex, and race/ethnicity.

As the overall model of predictors for the dependent variable that related to passing science, the resulting statistics (-2*Log Likelihood = 1027.470; χ^2 (9, N = 1018) = 176.924, $p < .001$) enabled me to reject this null hypothesis. Since those values were indeed significant, we can conclude that the predictor variables reduced the amount of unexplained variance that had been present in the base model. The fact that the -2*Log Likelihood statistic decreased by a value of 171.513 after including the four predictor variables also showed that those variables accounted for a significant amount of the unexplained variance in the overall regression model. Once the four predictor variables were entered into the regression model, the model also went from correctly classifying 66.6% of cases to correctly classifying 70.7% of cases. In other words, by using the four predictor variables the model could correctly predict whether a student passed his/her first science course 70.7% of the time (as opposed to only 66.6% percent of the time when those four variables were not in the model). In addition, the proportion of variance accounted for by the regression model (Cox & Snell R^2) went from 0.6% (without the four predictor variables) to 17.1% (when including the predictor variables). The fact that an additional 16.5% of the variance was accounted for by the full model clearly supports the use of these variables as a means for predicting success in first-year STEM majors' first science course. Once again, this result supported the literature that has emphasized the importance of high school preparation and that has found students' pre-college academic attributes useful for predicting college grades (Cole & Espinosa, 2008; Veenstra et al., 2008; Williford, 2009).

Null Hypothesis 7: Math SAT scores do not significantly predict success in the first science courses taken by first-year STEM majors when controlling for unweighted high school GPA, ALEKS math placement test scores, highest level of math taken in high school, year of entry, sex, and race/ethnicity.

Since the statistics for the overall model of predictors indicated that this predictor variable was not significant (*Wald* = 2.764, p = .096), we must fail to reject this null hypothesis. In addition to not being significant for the overall model, there were also additional results that indicated that this variable did not contribute significantly toward the effort to predict those who had passed their first science course. For example, the fact that the -2*Log Likelihood statistic only decreased by a value of 2.774 after including the math SAT variable demonstrated that it accounted for a relatively small amount of the unexplained variance in the overall regression model. The model also went from correctly classifying 70.9% of cases to correctly classifying 70.7% of cases. In other words, utilizing all of the predictor and control variables except for math SAT provided a correct prediction of passing science 70.9% of the time, while including the math SAT lowered that prediction rate to 70.7%. The proportion of variance accounted for by the model (Cox & Snell R^2) went from 16.8% (without math SAT score) to 17.1% (when including math SAT score), an increase of less than 1%. All of these data suggest that this variable had very little predictive power. Just like in the analysis for predicting those students who passed their first math course (research question two), this result goes against those studies which have found statistically significant correlations between SAT scores and first-year college grades/GPA (Burton & Ramist, 2001; Camara & Echternacht, 2000).

Null Hypothesis 8: Unweighted high school GPA does not significantly predict success in the first science courses taken by first-year STEM majors when controlling for math SAT scores, ALEKS math placement test scores, highest level of math taken in high school, year of entry, sex, and race/ethnicity.

Since the statistics for the overall model of predictors indicated that this predictor variable was significant (*Wald* = 96.333, $p < .001$), we must reject this null hypothesis. In addition to this variable's significance for the overall model, there were also additional results that indicated that this variable contributed significantly toward the effort to predict those who had passed their first science course. For example, the fact that the -2*Log Likelihood statistic decreased by a value of 117.13 after including the unweighted high school GPA variable demonstrated that it accounted for a significant amount of the unexplained variance in the overall regression model. The model also went from correctly classifying 66.7% of cases to correctly classifying 70.7% of cases. In other words, utilizing all of the predictor and control variables except for unweighted high school GPA provided a correct prediction of passing science 66.7% of the time, while including the unweighted high school GPA improved that prediction rate to 70.7%. The proportion of variance accounted for by the model (Cox & Snell R^2) went from 6.1% (without unweighted high school GPA) to 17.1% (when including unweighted high school GPA). The fact that an additional 11% of the variance was accounted for by this variable supports its use as a means for predicting those first-year STEM majors' who are likely to pass their first science course. It is also worth noting that while the overall model of predictors was significant, the unweighted high school GPA variable was the only individual predictor variable found to be significant for predicting success in science. This result once again supported those studies from the literature that have found high school grades to be the best predictor of the grades students will earn in college (Belfield & Crosta, 2012; Chase & Jacobs, 1989; Hoffman & Lowitzki, 2005; Nagashi & Slade, 2012; Thompson & Bolin, 2011; Williford, 2009; Zwick & Sklar, 2005).

Null Hypothesis 9: ALEKS math placement test scores do not significantly predict success in the first science courses taken by first-year STEM majors when controlling for math SAT scores, unweighted high school GPA, highest level of math taken in high school, year of entry, sex, and race/ethnicity.

Since the statistics for the overall model of predictors indicated that this predictor variable was not significant (*Wald* = 2.082, *p* = .149), we must fail to reject this null hypothesis. In addition to not being significant for the overall model, there were also additional results that indicated that this variable did not contribute significantly toward the effort to predict those who had passed their first science course. For example, the fact that the -2*Log Likelihood statistic only decreased by a value of 2.086 after including the variable for ALEKS score demonstrated that it accounted for a relatively small amount of the unexplained variance in the overall regression model. The model also went from correctly classifying 71.5% of cases to correctly classifying 70.7% of cases. In other words, utilizing all of the predictor and control variables except for ALEKS score provided a correct prediction of passing science 71.5% of the time, while including the ALEKS score lowered that prediction rate to 70.7%. The proportion of variance accounted for by the model (Cox & Snell R^2) went from 16.9% (without ALEKS score) to 17.1% (when including ALEKS score), an increase of less than 1%. All of these data suggest that this variable had very little predictive power. My literature review did not reveal anything related to placement results and grades in college science courses and the result is perhaps not that surprising given the fact that many of the students in my sample took science courses that had little or no quantitative content.

Null Hypothesis 10: The highest level of math taken in high school does not significantly predict success in the first science courses taken by first-year STEM majors when controlling for math SAT scores, unweighted high school GPA, ALEKS math placement test scores, year of entry, sex, and race/ethnicity.

Since the statistics for the overall model of predictors indicated that this predictor variable was not significant (*Wald* = 1.043, *p* = .307), we must fail to reject this null hypothesis. In addition to not being significant for the overall model, there were also additional results that indicated that this variable did not contribute significantly toward the effort to predict those who had passed their first science course. For example, the fact that the -2*Log Likelihood statistic only decreased by a value of 1.045 after including the variable for highest math taken in high school demonstrated that it accounted for a relatively small amount of the unexplained variance in the overall regression model. The model also went from correctly classifying 70.5% of cases to correctly classifying 70.7% of cases. In other words, utilizing all of the predictor and control variables except for highest math taken in high school provided a correct prediction of passing science 70.5% of the time, while including the highest math taken in high school only slightly improved that prediction rate to 70.7%. The proportion of variance accounted for by the model (Cox & Snell R^2) went from 17.0% (without highest level of math taken in high school) to 17.1% (when including highest level of math taken in high school), an increase of less than 1%. All of this suggests that this variable had very little predictive power. That was perhaps a bit surprising given the findings of studies like the one conducted by Tai et al. (2006), which found a link between high school math preparation and passing college chemistry. Chemistry, however, is a course that has significant math content and once again many of the students in my sample took science courses that were much less quantitative in nature.

Overall, when all four predictor variables were utilized together, the resulting models were statistically significant for predicting both those first-year STEM majors who passed their first math course and those who passed their first science course. As a result, those four predictor variables supported the literature related to college student retention and success that served as a guide for the development of my study. That literature has provided multiple examples that have found cognitive variables (such as course grades, GPA, and test scores) and other pre-college factors (such as the academic rigor of high school coursework) to be the most significant for predicting success in college (Cole & Espinosa, 2008; Engle & Tinto, 2008; Fuertes & Sedlacek, 1994; Levin & Wyckoff, 1991). There have also been several research studies that have found a statistical link between those cognitive measures and success in STEM disciplines (Nicholls et al., 2007; Sadler et al., 2012; Veenstra et al., 2008). One study in particular, conducted by Levin and Wyckoff (1991), measured the impact of 19 different variables (both cognitive and psychological) on the success of engineering majors at Penn State University. In doing so, they found that students' high school GPA, math SAT score, and whether or not they had taken Calculus in high school were the most correlated with students' success.

In examining the impact and predictive power of each of the four independent variables individually, unweighted high school GPA was found to be the best predictor and the only one that was significant for predicting success in both math and science. That finding definitely supported the theories developed from past research that high school GPA is the single best predictor of college success (Chase & Jacobs, 1989; Hoffman & Lowitzki, 2005; Kanoy et al., 1989; Williford, 2009; Zwick & Sklar, 2005). The other three predictor variables (math SAT, ALEKS score, and highest math taken in high school) were significant for predicting those who passed math but not significant for predicting those who passed science. Along those lines, the highest level of math taken in high school was the second best predictor (after unweighted high school GPA) for passing math. That finding supported the work of both Cox (2000) and Wheat et al. (1991) who both reported that the level of math preparation in high school was important to predicting success in college. A study conducted by Sadler et al. (2012) also found that students' success in high school math (especially Calculus) was a strong predictor of those students' likelihood to succeed in STEM majors. In support of those studies that have found SAT scores to be useful for predicting success in college, especially during the first year, the results of my study did indicate a statistical significance between math SAT scores and passing math (Burton & Ramist, 2001; Camara & Echternacht, 2000; Fuertes & Sedlacek, 1994; Patterson et al., 2012). However, the fact that the math SAT scores provided little in the way of improving the prediction of those students who passed math supported those studies that have critiqued the SAT for having little to no correlation with college success (Kanoy et al., 1989; Hiss & Franks, 2014). The study results also supported the research of Scott-Clayton (2012), who noted that placement test results are valid for predicting success in math by indicating that the inclusion of the ALEKS math placement test variable (within the overall model) was also significant. Similar to the SAT score, however, the Pseudo R^2 and Classification Model statistics pointed out that the ALEKS variable did not account for a very substantial amount of the model's unexplained variance.

As a whole, the results of my study should be of interest to those who are trying to identify factors that help predict college student success and those interested in researching students' early success as STEM majors. Given the nature of the institution from which the sample was selected and the students' demographics, the results might also prove useful for those who are attempting to either conduct research or implement strategies aimed at improving the retention rates for minority students in the STEM disciplines. Along those lines, the limits of my study's generalizability are discussed more thoroughly in the next section.

Limitations

As with all research, my study is limited by those factors that might impede the ability to generalize its results. First and foremost, the sample itself was limited by a number of factors. The students who comprised the sample population were all attending the same large, public, urban, Hispanic-serving institution in the southeastern U.S. As such, the results would be most generalizable to other first-year STEM majors who entered that same university, as opposed to those who might enroll at other institutions. Having said that, researchers and educators may be able to utilize and apply the results of my study at other institutions in that same region of the U.S. and at institutions with similar student demographics.

The sample was also limited to students who had entered college between 2010 and 2012. While that limitation prevented the use of data from prior years, it was an intentional decision aimed at acquiring the most recent and reliable data. Utilizing data from years prior to 2010 would have impacted the use of at least two of the four predictor variables. In the case of math SAT scores, going back too many years would have confounded the comparability among students since the test itself and the way it is scored have changed over time. In the case of the ALEKS math placement test score, the institution utilized for my study did not fully implement its use until 2010. As noted in both Chapters III and IV, even including students who entered in 2010 was problematic since the ALEKS scores were not consistently recorded in the student information system until 2011 (although the scores were recorded for certain first-year students in 2010 which enabled me to include some students who entered that year in my sample). Finally, the sample was also limited to those students who were pursuing one of the institution's 15 STEM majors. That was once again by design but does perhaps limit the ability to generalize the results to students who are pursuing non-STEM majors (even at the institution that was utilized for my study). With regard to each of these limitations to the sample, further research would need to be conducted to improve the generalizability of utilizing these traditional cognitive measures of academic preparation to predict college students' success in first-year math and science courses.

My research study was also limited by the fact that it focused on the use of four specific cognitive measures of students' academic preparation. For example, math SAT score was selected as one of those predictor variables despite the fact that students' overall or verbal SAT scores might have also proved to be useful in predicting success. As mentioned previously, 146 students were eventually excluded from the sample because they did not have an SAT score at all. In each of those cases, the students had completed the ACT test instead of the SAT. Many of the students who were included in the sample had both SAT and ACT test scores, but only the (math) SAT values were utilized for my study. The decision to utilize the ALEKS math placement test score as one of the predictor variables also limited both the sample (as mentioned in the previous paragraph) and the generalizability of the results. With regard to the sample, 184 students were removed because they did not have an ALEKS score. Even more limiting, only certain institutions utilize the ALEKS test as their math placement instrument. For those institutions who utilize a different math placement assessment, my study might serve to encourage a consideration of how their assessment correlates with predicting student success. For those institutions who do not utilize a math placement test at all, the use of that particular predictor variable would obviously not be an option. Three of the four predictor variables utilized for my study were also related to students' math aptitude which proved more useful for predicting students' who passed math than it did for predicting those who passed science. This decision to focus on math attributes even though science courses were included in the study limited my ability to develop models that might more accurately predict passing science.

The most significant limitation to the use of the highest level of math taken in high school related directly to an issue that was addressed in Chapter I. Even though the concepts and basic tenets of each level of mathematics are somewhat universal, the fact that the students (and students in general) attended several different high schools inherently raises questions about the comparability of their educational experiences. Even if the math content was exactly the same (which is unlikely), the method of instruction, access to resources, and even the location of the school would undoubtedly vary. The use of the students' unweighted high school GPA, on the other hand, was perhaps the least limiting of the four independent variables, particularly because the unweighted average was utilized instead of the students' weighted average. Past research has confirmed that point due to the fact that the weighted GPA is calculated after assigning additional points to students who have taken advanced courses, which involves methods that vary greatly among the various school districts in this country (Nagaishi & Slade, 2012). In addition, utilizing the unweighted GPA also created a more level measure of academic success for those students who were from outside the U.S. and for those who may not have had access to advanced courses while in high school. Having said that it is worth noting, regardless of the GPA in question, utilizing grades to measure an individual's competency or level of knowledge is a very subjective process.

Perhaps the most obvious limitation of my research lies in the fact that only cognitive factors were utilized as a means for predicting student success. The decision to focus on cognitive measures was rooted in both convenience (with regard to the applicability of the results, not my own process for collecting data) and the literature. By identifying suitable predictors among the information that institutions already know about their incoming students (as opposed to other variables that need to be measured or collected), the research results could be more easily implemented as a method for either placing STEM majors in appropriate courses or ensuring that at-risk students receive additional resources. Within the research, there is strong support for the fact that academic variables are the best predictors of academic success in college (Fuertes & Sedlacek, 1994). That support is even stronger when it comes to identifying the link between cognitive measures of academic knowledge and success in the STEM disciplines. Research studies conducted by Nicholls et al. (2007), Sadler et al. (2012), and Veenstra et al. (2008), for instance, all produced results that supported the fact that cognitive variables such as high school GPA and SAT scores were statistically relevant for predicting the academic success of STEM majors. As noted in Chapter I, the latter study found that 38% of the variance in the first-year GPA of engineering students was attributed to academic aspects of their pre-college preparation (Veenstra et al., 2008). Despite the extensive support for the predictive utility of cognitive variables, there have also been several research studies that have highlighted the predictive power of more psychosocial variables such as attitude, academic self-concept, and connection to campus (Cole & Espinosa, 2008; Kanoy et al., 1989; Wheat et al., 1991). While not the focus of my research study, Burton and Ramist (2001) have suggested that researchers consider using non-academic factors in conjunction with cognitive ones to reduce the limitations of the results and improve the validity of predicting academic success.

With regard to research design, there are always several decisions that could have been made from the onset to analyze the data differently (and perhaps in a less limiting fashion). For example, the statistical significance of the predictor variables was influenced by the decision to conduct a two-tailed test as opposed to a one-tailed test. That decision and the bulk of the other alternate research design decisions are addressed in the "Recommendations for Future Research" section that follows later in this chapter. The study results are not, however, limited by the lack of sample size. The use of over 1000 students provides enough confidence that the results could be useful for generalizing to the larger population (at least the population of students who pursue STEM majors at the institution utilized for my study). In an effort to confirm that notion, the GPower 3.1 software was utilized to run a post hoc analysis of the study's power. Utilizing various effect sizes (from small to large), the power of the logistic regression models were found to range from 0.91 to 0.98 (with 1 being the largest possible value for power). According to Long (1997), even though you need a larger sample size when you are utilizing multiple predictor variables, a sample size of greater than 500 is often large enough to produce regression models with meaningful (generalizable) results. With that being said, the large sample size could have been utilized to further improve the study's generalizability. By randomly dividing the sample into two and running the same statistical analyses twice (in an effort to replicate the results) I would have able to increase the estimates of replicability for my study.

Implications for Theory and Research

One of the motivations for my research study related to my desire to contribute to the existing literature on the factors that explain and contribute to college student success. A great deal of the theories and research regarding why college students struggle (and even dropout) have focused on various aspects of the higher education environment that have created unnecessary barriers to success. While there is certainly more that colleges and universities can do to remove such barriers, my study aimed to add to those theories that concentrate more on addressing the attributes of the students themselves, such as the previously mentioned research of Freeman et al., (2007). The fact that unweighted high school GPA was the variable that was most significant for predicting those students who passed math and science supports the theory that past academic performance (as measured by GPA) is linked to success in college. As Levin and Wyckoff (1991) noted in their research on persistence in undergraduate engineering programs, the best predictor of future behavior is past behavior. Combining that with the research on the critical nature of the first year of college, my study attempted to demonstrate the relationship between students' past academic performance and their likelihood for passing key first-year courses. The fact that the results of my study found a significant relationship between cognitive variables (such as high school GPA and the highest math course that students had taken in high school) and the ability to predict passing math and science courses further supports past research studies that have emphasized the importance of academic preparation. According to Freeman et al. (2007), for example, lack of academic preparation is one of the most significant variables for identifying those students who are at-risk of dropping out.

Along those lines, my study can be utilized to support those who believe that both college and university educators and those who research college student success need to adjust the way they think about student retention. According to Engle and Tinto (2008), theories surrounding why certain college students do not succeed have shifted toward an increased emphasis on those students' pre-college preparation. That shift is related, in part, to the findings of several studies that have demonstrated the ways that today's students are academically unprepared for the rigors they experience in college (Daley, 2010; Stratton et al., 2008; Tinto, 1993). Applying that concept to those students pursuing STEM degrees, the work of both Cole and Espinosa (2008) and Veenstra et al. (2008) found that pre-college characteristics, such as high school performance and especially math preparation, were the most significant variables for predicting those STEM students' success. Throughout most of the literature, college student success has typically been measured by students' college GPA. What sets my research study apart is that it focused on measuring students' success (as defined by passing with a grade of C or higher) in specific courses that are required by most STEM degrees. As such, the results of my study not only support past research studies but also offer opportunities for exploring new theories by providing an example of the relationship between STEM students' pre-college preparation (especially unweighted high school GPA and the highest math taken in high school) and their success in the required math and science curriculum.

Finally, while the bulk of the literature that has utilized cognitive measures and pre-college preparation to predict future success has centered on the use of high school GPA and standardized test scores (such as SAT and ACT), my study incorporated a predictor variable that was associated with the highest level of math course that each student had completed in high school. The decision to include that variable was motivated by research that has found math preparation to be a significant factor in predicting success in college math (Cox, 2000; Wheat et al., 1991). The work of Engle and Tinto (2008), in particular, highlighted the fact that taking a more rigorous schedule of math and science courses in high school has been linked to those students who graduate from college. Williford's (2009) research concurred with that assessment by citing several studies that have found that increasing the intensity of the high school curriculum was the best strategy for improving success in college (especially during the first year). By looking at the highest level of math that the students took in high school, my study has provided results that reinforce that link more specifically for those STEM majors who are required to take more rigorous math courses than their non-STEM counterparts. All of this also suggests that there is more work to be done with regard to how colleges and universities partner with high schools to influence those curriculum decisions. In addition, the fact that my study was conducted by utilizing students from a majority minority-serving institution means that the results can also serve to inform those theories and research studies that strive to explain the performance of minority students. According to Smith (1995), higher education institutions need to continue to address the lower retention rates of their underrepresented minority populations, especially those who are pursuing degrees in the STEM disciplines. Utilizing various elements of students' pre-college preparation might be the key to those efforts since high school performance has been shown to provide the best indication that minority students will succeed in science-related majors (Cole & Espinosa, 2008).

Implications for Practice

What, then, do the results of my study mean for those individuals who are interested in assisting first-year STEM majors and predicting their likelihood to succeed? In the words of Williford and Wadley (2008), our goal as educators is to do whatever we can to enable student success. Since we do not have much (or perhaps any) control of students' pre-college preparation, we should take more responsibility for utilizing the information we know about that preparation to assist them with avoiding potential pitfalls. Anyone who has worked in higher education in recent years is aware of the greater emphasis that has been placed on student retention (and timely graduation). On the one hand, those retention efforts are linked with the financial ramifications associated with students who leave. In addition to policies that threaten to reduce state and federal funding (for those institutions who receive it) to those institutions with lower retention and graduation rates, the students themselves pay tuition. According to Kim et al. (2010), in these times of economic difficulty institutions cannot afford the lost revenue that is associated with those students who decide to leave.

Financial considerations aside, our efforts to retain students should also be motivated by our desire to ensure that students have more opportunities later in life. That is not to say that opportunities do not exist for those who do not go to college, but there is evidence to support that college graduates have both higher employment rates and salaries than those who attended college but did not complete a degree. According to the National Center for Education Statistics (2012), individuals with an earned bachelor's degree are about 12% more likely to be employed and earn an average of 55% more than those who dropped out of college. If we truly are interested in helping our students benefit from those greater opportunities, we must continue trying to figure out why certain students succeed and why others choose to leave. While there are certainly several variables to consider, Johnson (2006) noted that there is a significant link between those students who leave college and poor grades. As mentioned previously, we also know from past research that the first year of college is critical to our efforts to retain students (Stovall, 2000; Thompson & Bolin, 2011). With those two points in mind, my research study was conducted to try and develop a model for predicting those STEM students who might be likely to receive poor grades in math and science during their first year. According to Jalomo (2000), students who experience academic success and develop a positive academic self-concept early on are more likely to persist. All the more reason to develop prediction models like the ones utilized in my study so that those students who are more likely to fail can receive additional support and resources or be placed into less rigorous courses at the beginning of their college careers.

Along those lines, my research study has definite implications for those institutional practices that are designed to predict students' likelihood for success. Those implications certainly connect well to any efforts aimed more specifically at predicting the success of STEM majors. They may also, however, stretch as far as helping institutions predict whether all first-year students are likely to pass math and science courses. In either case, colleges and universities that are interested (as they should be) in predicting student success should be advised to follow the lead of those research studies that have declared academic variables and pre-college factors to be the best predictors (Cole & Espinosa, 2008; Fuertes & Sedlacek, 1994; Sadler & Tai, 2001). More specifically, high school GPA has been found to be the single best predictor of college success, as was noted earlier in this chapter and supported by the results of my study (Chase & Jacobs, 1989; Hoffman & Lowitzki, 2005; Williford, 2009; Zwick & Sklar, 2005). While those research studies found that to be true with regard to predicting students' college GPA, my study found that it was also true for predicting success in specific courses (math and science).

It has also been noted previously that there are several research studies that have supported the use of SAT scores for the purposes of predicting students' future success (Burton & Ramist, 2001; Camara & Echternacht, 2000; Fuertes & Sedlacek, 1994; Patterson et al., 2012). On the other hand, there are also researchers who have found the use of the SAT to be ineffective in predicting students' aptitude for success (Kanoy et al., 1989; Hiss & Franks, 2014). The results of my study certainly supported those who have critiqued the use of SAT scores. In all fairness, even those studies that have found a correlation between SAT and student success have focused on predicting student's GPA at the end of the first year of college. As such, they have warned against utilizing that relationship to make broad decisions about students' long-term success or eventual graduation from college (Burton & Ramist, 2001; Camara & Echternacht, 2000). While my research study did focus on the students' first-year of college, it measured success using their grades in specific courses as opposed to overall GPA. While the math SAT score was found to be a significant variable (statistically speaking) for predicting those who passed math (but not science), it did not provide any improvement for those predictions over and above the use of the other three predictor variables that were utilized for my study. As such, math SAT score may not be the most useful cognitive factor to consider when attempting to predict first-year STEM majors' success in specific courses.

With regard to STEM majors, the results of my research study can also serve to influence higher education practices related to ensuring their success. As noted earlier, the research of Thompson and Bolin (2011) found that the highest rate of STEM dropouts occur during those students' first year in college. They also noted that those early departures are most likely linked with the fact that students are not prepared to succeed in the rigorous math and science courses that they are expected to take. Once again, my study found that students' unweighted high school GPA was significant in predicting those students who passed their first math course and those students who passed their first science course. As such, institutions should consider evaluating factors such as high school GPA as part of their efforts to support those students who have a desire to enter the STEM disciplines. Providing early opportunities for academic success is an important strategy for retaining those students both at the university level and within the STEM fields. According to Kessel and Linn (1996), many of those students who begin college as STEM majors eventually change their majors. That is particularly problematic given the increased pressure that institutions have been receiving to not only improve graduation rates but to also increase their number of STEM graduates (Sadler et al., 2012; Thompson & Bolin, 2011). Overall, the results of my study have implications for informing our ability to predict those students who are likely to succeed so that we can provide additional support for those who are more likely to struggle. As I have mentioned previously, that could involve adapting the curriculum, requiring students to take alternative or review courses, or providing an array of additional support and resources for our first-year STEM majors.

While the majority of these implications and recommendations have focused on what else colleges and universities can be doing for students once they enter college, I wanted to end this section by revisiting a statement I made previously about our lack of control over students' pre-college preparation. While that may be the reality, it does not diminish the fact that my results have reinforced those findings in the literature that have shown a link between students' pre-college preparation and their ability to achieve higher college GPAs and ultimately graduate. With that being said, all of the interventions and additional support services that have already been mentioned in this section may not be enough to fully ensure that colleges and universities are doing everything they can to support the graduation rates of STEM majors. By reaching out to and working with those responsible for designing the K-12 curriculum we can attempt to have a greater influence on those pre-college characteristics that have proven useful in predicting future academic success. With regard to STEM, that would include a greater emphasis on requiring students to complete additional or higher level mathematics courses if they express an interest in those areas. Since students are often placed in tracks that determine their math curriculum early on, those conversations need to extend beyond just the high school curriculum. As I mentioned in Chapter II, that is supported by past studies that have found that students who are struggling to keep up with the math curriculum as early as the eighth grade are far less likely to succeed as STEM majors when they get to college (Nicholls et al., 2007; Nicholls et al., 2010). Overall, the results of my study and several past studies suggest that doing more to influence students' earlier educational opportunities can have a significant impact on increasing the number of students who eventually graduate from college.

Recommendations for Future Research

As mentioned earlier in this chapter, the recommendations for future research can most certainly correspond with a number of decisions related to the research design of my study. Based on my results, conducting a one-tailed test of significance (as opposed to the two-tailed test that I utilized) would have definitely impacted that statistical significance of my study's predictor variables. In addition, the fact that three of the variables related to math aptitude could have been explored further by analyzing the level of interaction among those predictor variables. Even further, while the test for multicollinearity confirmed that the predictor variables were not problematically correlated with one another, the overlap in variance that was being accounted for by each variable could be measured and possibly explained further by first conducting a factor analysis. The specific independent and dependent variables that were selected for my study could also be adjusted for future studies. First and foremost, given that the cognitive measures that were selected were found to be significant for predicting success in math and science courses, future researchers may want to consider other cognitive factors or other combinations of factors that might also provide significance for predicting future success. Some of those are discussed throughout the rest of this section, but my recommendations begin with offering possible adaptations to the dependent variables that were utilized in my study.

For future studies, researchers might consider utilizing a continuous dependent variable such as STEM students' first-year GPA as opposed to a binary variable related to passing specific courses. While that might duplicate some of the existing literature (which has tended to utilize GPA as the dependent variable) it would provide an alternative means for quantifying the relationship between the predictor variables and student success. One suggestion that would ensure that the research being conducted is contributing original knowledge while still utilizing GPA as a dependent variable is to consider regression models that aim to predict a subset of the college GPA, such as GPA in college math courses, GPA in college science courses, or GPA in college math and science courses. Doing so would enable the researcher (and those who might utilize the results) to obtain a more clear understanding of how the cognitive predictor variables relate to success in math and science courses by measuring the grades that were earned and not (as my study did) simply predict whether or not the students passed with a grade of C or higher. Along those lines, future research studies might also utilize a grade higher than C as the measure of succeeding in a course. Doing so could provide more confidence that students had achieved a sufficient level of proficiency in the foundational math and science courses within the STEM curriculum.

For those future research studies that aim to build on the effort to predict success in college math courses, there are a number of considerations that would expand upon the results of my study. For example, since my study found unweighted high school GPA to be the most significant variable for predicting those who passed math, future research should consider running a logistic regression analysis that further isolates the significance of that variable. That could be accomplished by not utilizing the other three predictor variables (math SAT, ALEKS score, and highest math taken in high school) as control variables for that regression analysis so that those variables are not included in the model. Shifting attention to the next most significant variable from my study, future researchers might also consider incorporating the grades earned or how well students did in their highest level of high school math as opposed to just the level itself. Moreover, the fact that unweighted high school GPA and the highest level of math taken in high school were the two variables that improved the prediction models the most suggests that more research could be done utilizing just those two predictor variables (without including the other two predictors in the model).

Since my study did not reveal much in the way of predicting the students' success in science courses, there are definitely possibilities for future research studies utilizing that as the dependent variable. The fact that unweighted high school GPA was the only significant variable for predicting those who passed science suggests the need for further analyses that utilize only that variable. As mentioned in the last paragraph, one way to do that would be to leave out the other three predictor variables (by not including them at all in the prediction model). Another recommendation for future research involves considering what other variables might be better suited to predicting success in science. For example, since the math SAT was not a significant variable perhaps others measures of standardized testing such as the scores that students receive on the verbal subsection of the SAT or the science subsection of the ACT would serve as better predictors. As I gathered the information on my sample population's science courses I noticed that a significant number of those who had not passed had been enrolled in chemistry courses (typically General Chemistry I). Even before running the regression analyses, I wondered if my three math-related predictor variables might be more suitable for predicting success in those science courses that require more math aptitude (such as chemistry and physics) as opposed to all science courses. As such, future research studies could disaggregate science from math by developing alternative models for predicting success in the specific science courses that have traditionally high failure rates.

Finally, my research study and all of the recommendations for future research that I have offered thus far have focused on utilizing cognitive measures as the only predictor variables. As previously noted, those cognitive and academic factors have been proven time and time again to be the best predictors of future success. That does not mean, however, that they are the only factors that can prove significant for predicting students' success in college. Once researchers identify the best cognitive variables to include in prediction models, they should consider following the recommendation of Burton and Ramist (2001) and include both academic and non-academic variables. As they noted, combining such psychosocial variables as attitude toward academics, motivation to earn a degree, and academic self-concept with the cognitive variables related to students' academic preparation will improve the statistical validity of efforts to predict college students' academic success.

Summary and Conclusions

In reviewing the results of my quantitative research study, there are a number of conclusions that come to mind. First and foremost, the use of cognitive variables resulted in statistically significant models for predicting whether the students were successful in their math and science courses. More specifically, logistic regression models were conducted so that the -2*Log Likelihood statistic, full Chi-Square model, Pseudo R^2 (Cox & Snell R^2), and classification model could be utilized to test whether math SAT scores, unweighted high school GPA, ALEKS math placement test scores, and the highest level of math taken in high school were significant for predicting whether a sample of 1018 first-year STEM majors passed their first math and science courses in college. Those 1018 students had all entered the same large, public, urban, Hispanic-serving, research university from 2010 to 2012 and had indicated that they planned to major in one of that institution's 15 STEM majors.

In addition to the significance of the overall model of predictors, logistic regression analyses were also conducted to measure the significance of each of the predictor variables individually. All four predictors were found to be significant for predicting those who had passed their first math course (-2*Log Likelihood = 1210.015; χ^2 (9, N = 1018) = 196.687, p < .001). With that being said, the unweighted high school GPA and highest level of math taken in high school accounted for the greatest amount of variance in the model (in that order). As mentioned previously, the data that were presented in the classification model indicated that the use of those particular variables improved the ability to correctly predict if students passed math by up to as much as 15%. Conversely, the data also indicated that including the math SAT score and the ALEKS score did not result in any improvement in predicting those students who passed their first math course. With regard to the other dependent variable, only unweighted high school GPA was found to be significant for predicting those who had passed their first science course (-2*Log Likelihood = 1027.470; χ^2 (9, N = 1018) = 176.924, p < .001). That variable accounted for an additional 11% of the proportion of the variance for that model and improved the ability to correctly predict if students passed science by 4%.

As in all research, there were noted limitations to how the results can be generalized to the population of first-year STEM majors. As discussed, those limitations related to the sample itself (although not the size of the sample), the use of the specific cognitive predictor variables, and the fact that there were no non-cognitive variables utilized as predictors. The results of my study both confirmed and complemented the existing literature on college student retention, predicting student success, and the factors associated with success in the STEM disciplines. More specifically, the results confirmed those studies that have found academic variables and pre-college factors to be the best predictors for success in college (Cole & Espinosa, 2008; Fuertes & Sedlacek, 1994; Engle & Tinto, 2008; Fuertes & Sedlacek, 1994; Levin & Wyckoff, 1991), as well as those studies that have found high school GPA to be the single best predictor (Chase &

Jacobs, 1989; Hoffman & Lowitzki, 2005; Williford, 2009; Zwick & Sklar, 2005) and those that have found that the level of high school math preparation is important to efforts to predict success in college math (Cox, 2000; Sadler et al., 2012; Wheat et al., 1991).

In addition to the recommendations for future research that were provided in this chapter, my hope is that my study will also be utilized by researchers and practitioners alike for the purpose of assisting first-year STEM majors with their educational pursuits. In addition to the financial benefits of retaining students, efforts to predict their likelihood for success might be the key to building the confidence and academic self-concept that they need to persist and graduate. The models presented in my study can hopefully influence those policies that promote the necessity for predicting students' likelihood for success as well as those practices that provide additional resources for those students who are more likely to experience academic difficulty. As educators, we have a responsibility to reduce those barriers that impede student success. By utilizing the existing literature and conducting new research studies, we can work toward gaining a better understanding of what college students need to succeed. By applying that knowledge to those students who are pursuing the math and science-related disciplines we can also enhance our ability to support their first-year transition. In doing so, institutions can hopefully rely on their STEM majors' early success to increase the overall number of students (including minority students) who graduate with STEM degrees.

REFERENCES

ACT, Inc. (2010). *What works in student retention? Fourth national survey. Report for all colleges and universities* (2010). Retrieved January 17, 2013 from: http://search.proquest.com/docview/742868672?accountid=10901

Aragon, S. (2000). *Beyond access: Methods and models for increasing retention and learning among minority students. New Directions for Community Colleges, 28*(4)*. The Jossey-Bass higher education and adult education series.* San Francisco, CA: Jossey-Bass.

Armstrong, W.B. (2000). The association among student success in courses, placement test scores, student background data, and instructor grading practices. *Community College Journal of Research and Practice*, *24*(8), 681-695.

Belcheir, M., Michener, B., & Gray, N. (1998). Who stays? Who leaves? Results from a qualitative freshmen study. Research Report for Boise State University. Boise, ID. October, 1998.

Belfield, C. & Crosta, P.M. Teachers College, Community College Research Center (2012). *Predicting success in college: The importance of placement tests and high school transcripts.* (Working paper, No. 42). New York, NY.

Blanchard, W., & Mascetti, K. (2000, April). *Tracking down nonreturning students at an urban university: Method and results.* Paper presented at the Annual Meeting of the Alabama Association for Institutional Research, Gulf Shores, AL.

Burton, N.W., & Ramist, L. (2001). *Predicting success in college: SAT studies of classes graduating since 1980* (Research Report No. 2001-2). New York, NY: College Board.

Camara, W.J., & Echternacht, G. (2000). The SAT I and high school grades: Utility in predicting success in college. College Board Research Report RN-10. New York, NY.

Chase, C.I., & Jacobs, L. (1989). Predicting college success: The utility of high school achievement averages based on only “academic” courses. *College and University*, *64,* 403-408.

Cimetta, A.D., D’Agostino, J.V., & Levin, J.R. (2010). Can high school achievement tests serve to select college students? *Educational Measurement: Issues and Practice, 29*(2), 3-12.

Cole, K.C. (2001). Brain’s use of shortcuts can be a route to bias. In J. Ferrante & P. Brown (Eds.), *Social construction of race and ethnicity in the United States* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Cole, D., & Espinoza, A. (2008). Examining the academic success of Latino students in science, technology, engineering and mathematics (STEM) majors. *Journal of College Student Development*, *49*(4), 285-300.

Cox, W.W. (2000). Predicting the mathematical preparedness of first-year undergraduates for teaching and learning purposes. *International Journal of Mathematical Education in Science and Technology*, *31*(2), 227-248.

Daley, F. (2010). Why college students drop out and what we do about it. *College Quarterly, 13*(3), 1-5.

Davis, J.D., & Shih, J.C. (2007). Secondary options and post-secondary expectations: Standards-based mathematics programs and student achievement on college mathematics placement exams. *School Science and Mathematics*, *107*(8), 336-346.

Engle, J., & Tinto, V. (2008). Moving beyond access: College success for low-income, first-generation students. The Pell Institute for the Study of Opportunity in Higher Education. Washington, DC.

Freeman, J., Hall, E., & Bresciani, M. (2007). What leads students to have thoughts, talk to someone about, and take steps to leave their institution? *College Student Journal, 41*(4), 755-770.

Fuertes, J.N., & Sedlacek, W.E. (1994). Predicting the academic success of Hispanic college students using SAT scores. *College Student Journal, 28*, 350-352.

Glogowska, M., Young, P., & Lockyer, L. (2007). Should I go or should I stay?: A study of factors influencing students' decisions on early leaving. *Active Learning in Higher Education*, *8*(1), 63-77.

Haislett, J., & Hafer, A.A. (1990). Predicting success of engineering students during the freshman year. *Career Development Quarterly*, *39,* 86-95.

Harackiewicz, J.M., Barron, K.E., Tauer, J.M., & Elliot, A.J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology*, *94*(3), 562-575.

Hemmings, B., Grootenboer, P., & Kay, R. (2011). Predicting mathematics achievement: The influence of prior achievement and attitudes. *International Journal of Science & Mathematics Education*, *9*(3), 691-705.

Hiss, W.C., & Franks, V.W. (2014). Defining promise: Optional standardized testing policies in American college and university admissions. National Association for College Admission and Counseling. Arlington, VA.

Hoffman, J.L., & Lowitzki, K.E. (2005). Predicting college success with high school grades and test scores: Limitations for minority students. *The Review of Higher Education*, *28*(4), 455-474.

Hoyt, J.E., & Sorensen, C.T. (2001). High school preparation, placement testing, and college remediation. *Journal of Developmental Education*, *25*(2), 26-33.

Jalomo, R. (2000). Assessing minority student performance. *New Directions for Community Colleges,* 7-18. doi: 10.1002/cc.11201

Johnson, I. (2006). Analysis of stopout behavior at a public research university: The multi-spell discrete-time approach. *Research in Higher Education, 47*(8), 905-932.

Kanoy, K., Wester, J., & Latta, M. (1989). Predicting college success of freshmen using traditional, cognitive, and psychological measures. *Journal of Research and Development in Education, 22*(3), 65-70.

Kessel, C., & Linn, M.C. (1996). Grades or Scores: Predicting future college mathematics performance. *Educational Measurement: Issues and Practice*, *15*(4), 10-14.

Kim, E., Newton, F., Downey, R., & Benton, S. (2010). Personal factors impacting college student success: constructing college learning effectiveness inventory. *College Student Journal, 44*(1), 112-125.

Lehmann, W. (2007). “I just didn’t feel like I fit in”: The role of habitus in university dropout decisions. *Canadian Journal of Higher Education, 37*(2), 89-110.

Levin, J., & Wyckoff, J. (1991). Predicting persistence and success in baccalaureate engineering. *Education*, *111*(4), 461-468.

Long, J.S. (1997). *Regression models for categorical and limited dependent variables.* Thousand Oaks, CA: SAGE Publishing, Inc.

Martin, N., & Meyer, K. (2010). Efforts to improve undergraduate student retention rates at a Hispanic serving institution: Building collaborative relationships for the common good. *College and University, 85*(3), 40-49.

Mathiasen, R.E. (1984). Predicting college academic achievement: A research review. *College Student Journal, 18,* 380-386.

McNeil, K., Newman, I., & Fraas, J.W. (2012). *Designing general linear models to test research hypotheses.* Lanham, MD: University Press of America, Inc.

Mertler, C. A., & Vannatta, R. A. (2005). *Advanced and multivariate statistical procedures*. Glendale, CA: Pyrczak Publishing.

Nagaishi, C., & Slade, M.K. (2012). *Are weighted or unweighted high school grade point averages better indicators of college success?* Paper presented at the National Conference on Undergraduate Research, Ogden, Utah.

National Center for Education Statistics. (2012). *Digest of Education Statistics.* Retrieved June 16, 2014 from the National Center for Education Statistics Web site: http://nces.ed.gov/programs/digest/

Newman, I. (1989, October). *There is no such thing as multivariate analysis: All analyses are univariate*! Paper presented at the Midwestern Education Research Association, Chicago, IL.

Nicholls, G.M., Wolfe, H., Besterfield-Sacre, M., & Shuman, L.J. (2010). Predicting STEM degree outcomes based on eighth grade data and standard test scores. *Journal of Engineering Education*, *99*(3), 209-223.

Nicholls, G.M., Wolfe, H., & Besterfield-Sacre, M., Shuman, L.J., & Larpkiattaworn, S. (2007). A method for identifying variables for predicting STEM enrollment. *Journal of Engineering Education*, *96*(1), 33-44.

Noble, J.P., & Sawyer, R.L. (2004). Is high school GPA better than admission test scores for predicting academic success in college? *College and University*, *79*(4), 17-22.

Norusis, M.J. (1998). *SPSS 8.0: Guide to data analysis.* Upper Saddle River, NJ: Prentice Hall.

Otero, R., Rivas, O., & Rivera, R. (2007). Predicting persistence of Hispanic students in their 1st year of college. *Journal of Hispanic Higher Education, 6,* 163-173.

Patterson, B.F., Mattern, K.D., & Swerdzewski, P. (2012). Are the best scores the best scores for predicting college success? *Journal of College Admission*, *217*, 34-45.

Roth, J., Crans, G.G., & Carter, R.L. (2001). Effect of high school course-taking and grades on passing a college placement test. *High School Journal*, *84*(2), 72-87.

Sadler, P.M., Sonnert, G., Hazari, Z., & Tai, R. (2012). Stability and volatility of STEM career interest in high school: A gender study. *Science Education*, *96*(3), 411-427.

Sadler, P.M., & Tai, R.H. (2001). Success in introductory college physics: The role of high school preparation. *Science Education*, *85*(2), 111-136.

Schneider, M. (2010, October). *Finishing the first lap: The cost of first-year student attrition in America's four-year colleges and universities*. Retrieved April 20, 2013 from http://www.air.org/files/AIR_Schneider_Finishing_the_First_Lap_Oct101.pdf

Schugurensky, D. (2003). Higher education restructuring in the era of globalization. In R.F. Arnove & C.A. Torres (Eds.). *Comparative education: The dialectic of the global and the local* (2nd ed.). Lanham, MD: Rowman & Littlefield.

Scott-Clayton, J.E. Teachers College, Community College Research Center (2012). *Do high-stakes placement exams predict college success?* (Working paper No. 41). New York, NY.

Siegel, M.J. (2011). Reimagining the retention problem: Moving our thinking from end-product to by-product. *About Campus, 15:* 8-18. doi: 10.1002/abc.20043

Smith, T.Y. (1995). The retention status of underrepresented minority students: An analysis of survey results from sixty-seven U.S. colleges and universities. AIR 1995 Annual Forum Paper.

Stovall, M. (2000). Using success courses for promoting persistence and completion. *New Directions for Community Colleges,* 44-54. doi: 10.1002/cc.11204

Stratton, L., O'Toole, D., & Wetzel, J. (2008). A multinomial logit model of college stopout and dropout behavior. *Economics of Education Review, 27,* 319-331.

Tai, R.H., Ward, R., & Sadler, P.M. (2006). High school chemistry content background of introductory college chemistry students and its association with college chemistry grades. *Journal of Chemical Education*, *83*(11), 1703-1711.

Thompson, R., & Bolin, G. (2011). Indicators of success in STEM majors: A cohort study. *Journal of College Admission*, (212), 18-24.

Tinto, V. (1993). Leaving college: Rethinking the causes and cures of student attrition (2nd ed.). Chicago, IL: The University of Chicago Press.

Tinto, V. (2008). When access is not enough. *The Carnegie Foundation for the Advancement of Teaching, August 2008.*

Veenstra, C.P., Dey, E.L., & Herrin, G.D. (2008). Is modeling of freshman engineering success different from modeling of non-engineering success? *Journal of Engineering Education*, *97*(4), 467-479.

Watson, L. (2010). *Retention and graduation rates as performance indicators in 2-year and 4-year postsecondary institutions*. (Unpublished doctoral dissertation). University of New Orleans, New Orleans, LA.

Wheat, J., Tunnell, J., & Munday, R. (1991). Predicting success in college algebra: Student attitude and prior achievement. *College Student Journal*, *25,* 240-244.

Whelan, C. B. (2011). Helping first-year students help themselves. *The Chronicle of Higher Education, 57*(33), A56.

Williford, A.M. (2009). Secondary school course grades and success in college. *College and University*, *85*(1), 22-33.

Williford, A.M. & Wadley, J. (2008). How institutional research can create and synthesize retention and attrition information. *Association for Institutional Research, 108,* 1-24.

Woosley, S., Slabaugh, K., Sadler, A., & Mason, G. (2005). The mystery of stop-outs: Do commitment and intentions predict reenrollment? *NASPA Journal, 42*(2), 188-201.

Young, A., Johnson, G., Hawthorne, M., & Pugh, J. (2011). Cultural predictors of academic motivation and achievement: A self-deterministic approach. *College Student Journal, 45*(1), 151-163.

Zwick, R., & Sklar, J.C. (2005). Predicting college grades and degree completion using high school grades and SAT scores: The role of student ethnicity and first language. *American Educational Research Journal*, *42*(3), 439-464.

VITA

CHARLES ANDREWS

Born, Miami, Florida

1994	B.S., Mathematics Education Florida International University Miami, Florida
1994 – 1995	Coordinator, Campus Life & Orientation Florida International University Miami, Florida
1997	M.A., College Student Personnel Bowling Green State University Bowling Green, Ohio
1998 – 1999	Academic Advisor, College of Education Bowling Green State University Bowling Green, Ohio
1999 – 2000	Coordinator, Academic Support Services Florida International University Miami, Florida
2000 – 2003	Associate Director, Orientation & Commuter Services Florida International University Miami, Florida
2003 – 2008	Director, Campus Life & Orientation Florida International University Miami, Florida
2008 – 2014	Director, Academic Advising & First-Year Programs Florida International University Miami, Florida
2010 – 2014	Doctoral Candidate, Higher Education Florida International University Miami, Florida