PREDICTING SUCCESS: AN EXAMINATION OF THE PREDICTIVE VALIDITY OF A MEASURE OF MOTIVATIONAL-DEVELOPMENTAL DIMENSIONS IN COLLEGE ADMISSIONS

A Dissertation
Submitted to
the Temple University Graduate Board

In Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF EDUCATION

---

by
Joseph H. Paris
May 2018

Examining Committee Members:

Dr. Joseph DuCette, Advisory Chair, Department of Psychological Studies in Education
Dr. Gregory Anderson, Department of Policy, Organizational, and Leadership Studies
Dr. Laura Pendergast, Department of Psychological Studies in Education
Dr. James Earl Davis, Department of Policy, Organizational, and Leadership Studies

ABSTRACT

Although many colleges and universities use a wide range of criteria to evaluate and select admissions applicants, much of the variance in college student success remains unexplained. Thus, success in college, as defined by academic performance and student retention, may be related to other variables or combinations of variables beyond those traditionally used in college admissions (high school grade point average and standardized test scores).

The current study investigated the predictive validity of a measure of motivational-developmental dimensions as a predictor of the academic achievement and persistence of college students as measured by cumulative undergraduate grade point average and retention. These dimensions are based on social-cognitive (self-concept, self-set goals, causal attributions, and coping strategies) and developmental-constructivist (self-awareness and self-authorship) perspectives. Motivational-developmental constructs are under-explored in terms of the predictive potential derived from their use in evaluating admission applicants' ability to succeed and persevere despite the academic and social challenges presented by postsecondary participation. Therefore, the current study aimed to generate new understandings to benefit the participating institution and other institutions of higher education that seek new methodologies for evaluating and selecting college admission applicants.

This dissertation describes two studies conducted at a large, urban public university located in the Northeastern United States. Participants included 10,149 undergraduate students who enrolled as first-time freshmen for the Fall 2015 (Study 1) and Fall 2016 (Study 2) semesters. Prior to matriculation, participants applied for admission using one of two methods: standard admissions or test-optional admissions. Standard admission applicants submitted standardized test scores (e.g., SAT) whereas test-optional applicants responded to four short-answer essay questions, each of which measured a subset of the motivational-developmental dimensions examined in the current study. Trained readers evaluated the

essays to produce a “test-optional essay rating score,” which served as the primary predictor variable in the current study. Quantitative analyses were conducted to investigate the predictive validity of the “test-optional essay rating score” and its relationship to cumulative undergraduate grade point average and retention, which served as the outcome variables in the current study.

The results revealed statistically significant group differences between test-optional applicants and standard applicants. Test-optional admission applicants are more likely to be female, of lower socioeconomic status, and ethnic minorities as compared to standard admission applicants. Given these group differences, Pearson product-moment correlation coefficients were computed to determine whether the test-optional essay rating score differentially predicted success across racial and gender subgroups. There was inconclusive evidence regarding whether the test-optional essay rating score differentially predicts cumulative undergraduate grade point average and retention across student subgroups.

The results revealed a weak correlation between the test-optional essay rating score and cumulative undergraduate grade point average (Study 1: $r = .11$, $p < .01$; Study 2: $r = .07$, $p < .05$) and retention (Study 1: $r = .08$, $p < .05$; Study 2: $r = .10$, $p < .01$), particularly in comparison to the relationship between these outcome variables and the criteria most commonly considered in college admissions (high school grade point average, SAT Verbal, SAT Quantitative, and SAT Writing). Despite these findings, the test-optional essay rating score contributed nominal value ($R^2 = .07$) in predicting academic achievement and persistence beyond the explanation provided by traditional admissions criteria. Additionally, a ROC analysis determined that the test-optional essay rating score does not predict student retention in a way that is meaningfully different than chance and therefore is not an accurate binary classifier of retention.

Further research should investigate the validity of other motivational-developmental dimensions and the fidelity of other methods for measuring them in an attempt to account for a greater proportion of variance in college student success.

DEDICATION

This dissertation is dedicated to my daughter, Everett. Your enduring joy, kindness, and love sustained me each day during my “big project.” May all of your dreams be fulfilled, always.

ACKNOWLEDGEMENTS

This dissertation would not be possible without the immeasurable contributions of the members of my dissertation committee to whom I extend my sincere and deepest gratitude. Dr. Joseph DuCette, I am eternally grateful for your mentorship, encouragement, and friendship. Without you, this accomplishment would not be a reality. Dr. Laura Pendergast, thank you for your invaluable feedback and expertise. Your insight and dedication helped me more fully recognize the impact and meaningfulness of my work. For this, I am most grateful. Dr. Gregory Anderson, thank you for serving as a perpetual source of inspiration and advocacy for my work. I am most appreciative of your dedication to my professional and academic success. Dr. James Davis, from the moment I arrived at Temple, you have supported and guided me in all of my academic and professional endeavors. I would not be where I am today if it were not for you.

Many others made instrumental contributions to this study. I extend my sincere appreciation to Dr. Jim Degnan, Dr. Avi Kaplan, and Dr. Jodi Levine Laufgraben for instilling me with a depth of knowledge regarding the background and institutional importance of this study. Your commitment to data-driven innovation and models of student success inspired me throughout this project. Thanks to my team members and colleagues in the College of Education for ensuring students have every opportunity to capitalize on their promise and ability.

To my parents, my first motivators and educators, thanks for always reminding me of my ability and helping me reach my fullest potential. I love you both beyond words.

To Aften, my wife, editor, and unwavering source of support and love, thanks for always being by my side each step of the way during this academic and personal journey. Words cannot express how deeply I love you and appreciate all you have done to help me reach this milestone.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Postsecondary education is a primary engine of social mobility in the United States (Baum, Ma, & Payea, 2013). Research in the field of higher education has identified many positive economic and noneconomic benefits of postsecondary participation and completion (Haveman & Smeeding, 2006). Admission to higher education serves as a perceived and actual gateway to success in the American economy and society. Given the prominence of postsecondary education as an economic and societal imperative, the process and methodology by which colleges and universities select college applicants for admission have been the subject of significant public debate and scholarly inquiry (Camara & Kimmel, 2005).

Haveman and Smeeding (2006) suggest, “Most Americans expect the nation’s colleges and universities to promote the goal of social mobility to make it possible for anyone with ability and motivation to succeed” (p. 125). However, individual differences in ability and motivation influence the performance and success of college students (Komarraju, Karau, & Schmeck, 2009). Therefore, there is applied interest in examining the criteria institutions of higher education use to evaluate the ability and motivation of college admission applicants (Atkinson & Geiser, 2009). The development of reliable and valid predictors of college student success is necessary to expand college participation, student achievement, and access to and equity in higher education, especially for underrepresented student populations (Sackett, Schmitt, Ellingson, & Kabin, 2001; Thomas, Kuncel, & Credé, 2007). To this end, there is the need to identify and validate new variables that consistently and accurately measure a broader range of attributes related to college student success (Camara & Kimmel, 2005).

Statement of the Problem

The primary objective of college admissions is the identification and selection of applicants who are most likely to achieve success in college (National Association for College Admissions Counseling, 2016). Institutions of higher education traditionally utilize a combination of variables to evaluate the ability of admission applicants including, but not limited to, high school grade point average and standardized admissions tests such as the SAT or ACT. Institutions may also assess applicants' motivation to succeed in college by measuring non-cognitive predictors including demographic information and attitudinal, behavioral, and personality constructs, primarily derived from psychological theories (Allen, Robbins, & Sawyer, 2010). These "nontraditional" variables often include biological and demographic data, psychosocial factors, and personality traits (Allen et al., 2010).

An extensive body of literature has investigated the efficacy of admission models and the validity of the criteria of which they are comprised (NACAC, 2016). Although there is a range of criteria traditionally considered in college admissions, the most accurate statistical models account for approximately 25-30% of the variance in college outcomes such as first-year grade point average (Atkinson & Geiser, 2009; Tross, Harper, Osher, & Kneidinger, 2000; Wolfe & Johnson, 1995). Much of this variance remains unexplained by traditional admission criteria.

A subtext of the debate regarding how institutions make admission decisions is a concern for fairness and the equitable distribution of educational opportunities (Camara & Kimmel, 2005). An increasing number of institutions of higher education have introduced new models for evaluating the potential of college admission applicants in response to research that has identified limitations of traditional admissions criteria, particularly the differential prediction of standardized tests across the racial, gender, and socioeconomic subgroups of college admission applicants. The relationship between definitions of success and institutional values has also contributed to the development of new admission models and more specifically, the emergence of the test-optional admission policies.

The test-optional movement in the United States primarily emerged in response to criticism of standardized admissions tests as inadequate and potentially biased measures of postsecondary promise (Atkinson & Geiser, 2009; Belasco, Rosinger, & Hearn, 2015). This criticism draws from research that suggests socioeconomic status has a strong positive linear relation to standardized test scores (Buckley, Letukas, & Wildavsky, 2018; Geiser & Studley, 2002; Soares, 2012; Syverson, 2007; Zwick, 2002). Additional research has identified predictive biases of standardized tests across ethnic and gender groups (Buckley et al., 2018; Soares, 2012; Zwick, 2002). Despite these findings, proponents of the use of standardized tests in college admissions cite the need to control for the variability of high school grading standards and the availability and quality of postsecondary preparation (Atkinson & Geiser, 2009; Buckley et al., 2018; Burton & Ramist, 2001; Syverson, 2007; Zwick, 2002). In response to these conditions, there is need at the national and institutional levels to validate new variables that consistently and accurately predict college student success in order to attain the broader goals of higher education (Burton & Ramist, 2001).

The emergence of test-optional admission policies has reduced the reliance on nationally-normed standardized tests and, at some institutions, has placed greater emphasis on the non-cognitive attributes of college admission applicants. Despite a growing body of research that has examined the predictive validity of non-cognitive variables in college admissions, findings are inconclusive as to the incremental contributions of psychosocial, attitudinal, and motivational factors in explaining the variance in college outcomes (Allen et al., 2010; Komarraju et al., 2009;

Robbins et al., 2004; Thomas et al., 2007). Therefore, there remains a need for clear and consistent evidence regarding the extent to which motivational variables relate to college student success (Komarraju et al., 2009).

Purpose of the Study

The purpose of the current study was to address gaps in understanding the extent to which motivational-developmental dimensions predict college student success. Additionally, the current study responded to the need to identify reliable and valid variables that incrementally predict college outcomes beyond the variance accounted for by the criteria traditionally used in college admissions such as high school grade point average and standardized test scores. A central objective of the current study was to investigate the relationship between the scores on a measure of preadmission motivational-developmental dimensions and subsequent student success in college as defined by cumulative undergraduate grade point average and retention.

Research Questions

As many questions regarding the predictive validity of motivational-developmental dimensions remain unanswered, the current study addressed the following primary research questions:

1. Is the test-optional essay rating score a significant predictor of college student success, as measured by cumulative undergraduate grade point average and retention?
    a. Is the test-optional essay rating score incrementally predictive of cumulative undergraduate grade point average and retention given the variance accounted for by traditional admission criteria (high school grade point average and SAT scores)?

b. At what test-optional essay rating cut score are test-optional admission applicants predicted to persist in college?
c. Which motivational-developmental dimensions, if any, contribute most to the prediction of cumulative undergraduate grade point average and retention?
2. Does the test-optional essay rating score differentially predict cumulative undergraduate grade point average and retention across racial and gender subgroups?

Additionally, the current study addressed the following secondary research questions:

1. On average, do test-optional admission applicants attain a greater/lesser cumulative undergraduate grade point average than what is predicted by their SAT scores?
2. Are the New Student Questionnaire (NSQ) items that measure student motivation significant predictors of college student success, as defined by cumulative undergraduate grade point average and retention?

Definitions of Key Terms

The following definitions ensure uniformity and understanding of the terminology used throughout the current study. I developed all definitions not accompanied by a citation. Chapters 2 and 3 include comprehensive definitions of the predictor and outcome variables examined in the current study.

<u>ACT</u>: Formerly known as American College Testing, the ACT is a standardized, college readiness test widely utilized in college admissions. The ACT provides information for understanding students' academic strengths and weaknesses and insights about their current interests and postsecondary aspirations (Condition of College and Career Readiness, 2016).

<u>Big Five Personality Traits</u>: The Big Five personality traits include neuroticism, extraversion, openness, agreeableness, and conscientiousness. These traits relate to a wide range of behaviors

including academic achievement and job performance. These traits form a framework for examining the relationship between personality, motivation, and achievement (Komarraju et al., 2009; Trapmann, Hell, Hirn, & Schuler, 2007; Tross et al., 2000; Wolfe & Johnson, 1995).

Cognitive Assessment: According to the *Standards for Educational and Psychological Testing* (2004), cognitive assessment is the process of systematically collecting test scores and related data to make judgements about an individual's ability to perform various mental activities involved in the processing, acquisition, retention, conceptualization, and organization of sensory, perceptual, verbal, spatial, and psychomotor information (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014).

Cognitive Variable: Cognitive variables generally refer to measures of general intelligence such as IQ tests. Several variables traditionally used for college admission selection decision-making such as reasoning tests (SAT and ACT), high-school grade point average, and high school course grades measure general cognitive and reasoning abilities, and may assess subject-specific knowledge and skills (Camara & Kimmel, 2005). According to Noftle and Robbins (2007), "Previous research (e.g., Rohde & Thompson, 2007) suggests that the SAT verbal test may be more strongly related to crystallized intelligence because of its vocabulary-related content, whereas the SAT math test may be more strongly related to fluid intelligence because of its numerical and spatial reasoning related" (p. 127). Frey and Detterman (2004) found that the SAT correlated .82 with a measure of "g" (or general intelligence).

Content-Related Validity Evidence: According to the *Standards for Educational and Psychological Testing* (2004), content-related validity evidence is "based on test content that supports the intended interpretation of test scores for a given purpose. Such evidence may address issues such as the fidelity of test content to performance in the domain in question and the degree to which test content representatively samples a domain…" (AERA, APA, & NCME, 2014, p. 218).

Convergent Evidence: Convergent evidence is based on the relationship between test scores and other measures of the same related construct (AERA, APA, & NCME, 2014).

Criterion Variable: See outcome variable.

Differential Prediction: See predictive bias.

Differential Validity: As described by Linn (1978), differential validity refers to differences in the magnitude of the correlation coefficients for different groups of test-takers.

Enrolled Students: Enrolled students are registered for one or more credits during the semester for which they were offered admission. Enrolled students may be registered in subsequent semesters examined in the current study.

Expected Family Contribution: Expected Family Contribution (EFC) is a measure of family financial strength and is calculated according to a formula established by federal law. This formula considers taxed and untaxed income, assets, and benefits such as unemployment or Social Security. Also considered are family size and the number of family members who will attend college during a given academic year (Expected Family Contribution [EFC], 2017).

First-time Freshmen: First-time freshmen are students who have not previously enrolled in postsecondary education prior to their admission to a college or university.

Institution: The terms "college" and "university" are used interchangeably to refer to postsecondary educational institutions. I made finer distinctions between these terms wherever necessary or used the term "institution of higher education" to capture the broad range of postsecondary institution types. The term “participating institution” refers to the institution at which the current study was conducted.

Motivational-Developmental Dimensions: Motivational-developmental dimensions are non-cognitive, behavioral, attitudinal, and personality constructs, primarily derived from psychological theories (Allen et al., 2010). These “nontraditional” variables relate to psychosocial factors and personality traits such as self-efficacy, self-concept, and achievement motivation (Allen et al., 2010). Chapter 2 includes expanded definitions of the motivational-developmental dimensions examined in the current study.

Non-Cognitive Variable: According to Sedlacek (2004), the term non-cognitive refers to “variables relating to adjustment, motivation and perception, rather than the traditional verbal and quantitative areas typically measured by standardized tests” (p. 36).

Outcome Variable: Outcome variables refer to the presumed effect in a non-experimental study. Researchers measure outcome variables to determine how different experiences and/or inherent characteristics affect outcomes (Mertens, 2014). The outcome variable is a measure of a criterion and used to evaluate the validity and utility of a predictor variable (Camara & Kimmel, 2005).

Pell Grant: The Federal Pell Grant Program provides need-based grants to low-income undergraduate and certain post-baccalaureate students to promote access to postsecondary education. Grant amounts are dependent on the student’s Expected Family Contribution (EFC, see above); the cost of attendance (as determined by the institution); the student’s enrollment status (full-time or part-time); and whether the student attends for a full academic year or less (Pell Grant, 2018).

Persistence: Persistence refers to students maintaining or completing their enrollment at any college or university in the postsecondary system (Hagedorn, 2005).

Postsecondary Education: The terms “postsecondary education” and “higher education” are used interchangeably to represent education levels beyond high school.

Postsecondary-Situational Factors: Postsecondary-situational factors refer to non-cognitive variables that may be predictive of college student success and persistence. These variables may include level of financial aid received, employment status, family obligations, and out-of-state enrollment (Allen et al., 2010).

Predictive Bias: Predictive bias refers to the systematic under- or over-prediction of criterion performance for people belonging to groups differentiated by characteristics not relevant to the criterion performance (AERA, APA, & NCME, 2014).

Predictive Validity: Predictive validity refers to evidence that indicates how accurately data collected at one time can predict criterion scores obtained at a later time (AERA, APA, & NCME, 2014). College admission validity studies involve research examining the factors that are most relevant in predicting college performance, with the goal of identifying students who will be successful or who are at-risk (NACAC, 2016).

Predictor Variable: Predictor variables refer to presumed “cause” in a non-experimental study. Predictor variables refer to independent variables in studies in which the researcher is interested in the effect of differences of an inherent characteristic (Mertens, 2014).

Retention: Retention focuses on the goal of an institution to keep students enrolled from one term of study to the next until completion of a degree and/or credential (Hagedorn, 2005). Tinto’s (1975) landmark student integration model theorizes that students who socially integrate into the campus community increase their commitment to the institution and are more likely to graduate. A variety of factors are related to college student retention including academic preparation, academic engagement, social engagement, financing college, and demographic characteristics.

SAT: Formerly the “Scholastic Aptitude Test”, the SAT is a standardized test used widely in college admissions to assess student reasoning based on knowledge and skills developed by the students in their course work (SAT, 2016). The SAT is also used to “gauge students' general analytic ability, as distinct from their mastery of specific subject matter, and thereby to predict performance in college" (Atkinson & Geiser, 2009, p. 666).

Socioeconomic Status: As defined by the American Psychological Association, socioeconomic status is the social standing or class of an individual or group. Socioeconomic status is the stratification of a population as measured as a combination of education, income, and occupation. Examinations of socioeconomic status often reveal inequities in access to resources as well as issues related to privilege, power, and control (American Psychological Association, 2007).

Test-Optional Admissions: Test-optional admission policies provide applicants the choice whether to submit standardized test scores as part of an application for admission. Institutions with test-optional admission policies may consider standardized test scores in making admission decisions for applicants for whom scores are available.

Traditional Admission Predictors: Traditional admission predictors refer to the criteria most widely considered by postsecondary institutions to evaluate college admission applicants and predict college student success. Traditional admission criteria include high school grades, high school grade point average, class rank, and standardized test (i.e., ACT or SAT) scores. “Traditional” is not an empirical description, but a normative standard against which other types of admission criteria are frequently considered lesser approximations.

Validity: The *Standards for Educational and Psychological Testing* note that "validity refers to the degree to which evidence and theory support the specific interpretation of test scores for the proposed uses of tests" (AERA, APA, & NCME, 2014, p. 11). If a researcher intends to use multiple interpretations of a test score for different purposes, validity evidence for each interpretation is required.

Validity Coefficient: According to Salkind (2007), "The validity coefficient is a statistical index used to report evidence of validity for intended interpretations of test scores and defined as the magnitude of the correlation between test scores and a criterion variable (i.e., a measure representing a theoretical component of the intended meaning of the test)" (p. 2).

CHAPTER 2

LITERATURE REVIEW

The criteria by which institutions evaluate admission applicants have a prominent role in influencing educational opportunities. Therefore, these variables are of significant interest to the public and have been the subject of considerable empirical investigation (Sackett, Kuncel, Arneson, Cooper, & Waters, 2009). Research has focused on the “factors that are most relevant in predicting college performance, with the goal of identifying students who will be successful or who are at-risk” (NACAC, 2016, p. 4).

A review of the literature revealed five general domains of the factors most frequently utilized in college admissions. These domains include cognitive and non-cognitive academic measures, temperament and personality traits, personal qualities and experiences, applicant characteristics and demographics, and other factors. Despite the current understandings of these domains, existing literature is limited in scope, having focused predominately on cognitive and non-cognitive academic measures. Given the aforementioned unexplained variance in college outcomes, there remains a need for clear and consistent findings regarding the validity of additional predictors that offer utility in college admissions by contributing additional predictive power to existing models of admission decision-making (Atkinson & Geiser, 2009; Camara & Kimmel, 2005). Moreover, scholars (e.g., Camara & Kimmel, 2005; Willingham, 1985) have emphasized the need to broaden the definition of college student success in order to evaluate new predictors.

Definitions of College Student Success

The development and identification of a broader range of predictors used in college admissions requires an understanding of the multidimensional nature of college student success

(Burton & Ramist, 2001; Camara & Kimmel, 2005). Definitions of student success vary widely among institutions and educational stakeholders (Davenport, Martinez-Saenz, & Rhine, 2012). Admission criteria derive their importance and relevance as accurate predictors of the outcomes of greatest interest and value to postsecondary institutions (Camara & Kimmel, 2005). Frequently, success in college refers to academic performance as measured by course grades and grade point average (Atkinson & Geiser, 2009; Burton & Ramist, 2001; Camara & Echternacht, 2000; Camara & Kimmel, 2005; NACAC, 2016). However, grade point average is limited in capturing the variance associated with the distinct dimensions of educational performance (Camara & Kimmel, 2005).

A review of the literature revealed that freshmen and cumulative undergraduate grade point average are the most frequent criterion variables examined in predictive validity research (NACAC, 2016). Grade point average and course grades principally measure general cognitive performance and therefore are limited in serving as a criterion variable for many other factors related to college success such as retention and degree completion (Camara & Kimmel, 2005). Grade point average and course grades may reflect motivation and other psychosocial factors. However, the variability of grading standards and practices across courses, academic disciplines, instructors, and institutions limits the quantification of the multidimensionality of college success (Camara, Kimmel, Scheuneman, & Sawtell, 2003). Camara and Kimmel (2005) suggest that new measures of college student success have their greatest promise among performance dimensions such as creativity, leadership, persistence, and citizenship rather than academic outcomes such as freshman grade point average.

In addition to the narrow definition of college student success, predictive models infrequently consider long-term indicators of college student success such as post-college employment and earnings, graduate school acceptance, and civic participation (Burton & Ramist, 2001; Willingham, 1985). Burton and Ramist (2001) suggest, “Although most predictive validity studies use first-year grades as a proxy for success in college, long-term success in college is an equally important criterion to many colleges and universities” (p. 1). A study by the National Association for College Admissions Counseling (NACAC) found that many prior admission validity studies “are limited in defining ‘success’ solely as a function of the first-year GPA, largely due to the ease and timeliness of the data availability, rather than a narrow view of the meaning of success” (NACAC, 2016, p. 5). Long-term predictions are also important in the context of non-traditional students as ethnic minority students and other underrepresented student populations may take longer to successfully transition to their institutional environment (Neville, Heppner, & Wang, 1997). Therefore, non-traditional students may not demonstrate their true academic abilities until later in college. Given these findings and the gaps in the literature, there is a need to validate new admission variables that predict this traditional measure (i.e., cumulative undergraduate grade point average) and reflect the broader definition of college success.

Predictive Validity of Traditional Criteria in College Admissions

The evaluative methodologies employed by institutions have traditionally emphasized a range of cognitive factors such as standardized test scores and other variables with cognitive and non-cognitive dimensions such as high school course grades, high school grade point average, and class rank (Burton & Ramist, 2001). These variables measure cognitive and reasoning abilities and therefore may be used to assess subject-specific knowledge and skills (Camara & Kimmel, 2005). Therefore, colleges and universities traditionally use these factors to evaluate the academic preparedness of admission applicants.

Since the 1970s, institutions have used standardized tests as an efficient mechanism for qualifying and selecting an increasing number of admission applicants (Camara & Kimmel, 2005). Standardized test scores continue to serve as predominant criteria for the evaluation and selection of college admission applicants. Despite the emergence of test-optional admission policies, a recent NACAC survey revealed that 78% of four-year institutions require standardized tests as part of an admission application (NACAC, 2016). Additionally, the SAT was administered to a record 1.6 million high school graduates from the class of 2016 (SAT, 2016).

The College Board's past and present descriptions of the SAT assert that it measures students' general analytic and reasoning ability as distinct from their intelligence and subject matter mastery, and thereby predicts academic performance in college (Atkinson & Geiser, 2009). Despite this intended distinction, some researchers have identified strong correlations (r = .82) between SAT scores and general intelligence (Frey & Detterman, 2004).

An extensive and growing body of research has examined the validity of SAT scores in predicting college performance (e.g., Burton & Ramist, 2001; Hezlett et al., 2001; Morgan, 1989; Wilson, 1983)[1]. Similarly, previous research has investigated the predictive validity of other traditional criteria frequently used in college admission decision-making such as high school grades and grade point average. An examination of large-scale predictive validity studies and meta-analyses revealed several consistent findings regarding the correlation between traditional variables and college performance (e.g., Burton & Ramist, 2001; Camara, & Echternacht, 2000;

---

[1] See Burton & Ramist (2001) and Morgan (1989) for summaries of predictive validity studies conducted since 1976. Hezlett et al. (2001) conducted a meta-analysis of approximately 3,000 SAT validity studies involving over 1 million students. Kobrin and Patterson (2011) suggest the number of past predictive validity studies related to college admissions number in the thousands.

Kobrin & Patterson, 2011; Morgan, 1989; Willingham, Lewis, Morgan, & Ramist, 1990; Zwick, 2002). For example, research has identified that the correlation of standardized test (ACT and SAT) scores with first-year college grade point average is about .4 on average (Zwick, 2002). As reported in the literature, this validity coefficient is slightly lower than the average correlation between high school grade point average and first-year college grade point average. Studies have found that the combination of standardized test scores and high school grade point average as predictors produces average correlations of .5 with first-year college grade point average (Zwick, 2002).

Burton and Ramist (2001) conducted a meta-analysis to expand upon a previous review of cumulative college grade point average of classes graduating college between 1930 and 1980 (Wilson, 1983). The researchers examined classes graduating between 1980 and the mid-1990s to identify the qualities of students that have proven to predict success in college. Their findings suggest that academic preadmission measures (e.g., high school grade point average, course grades, standardized test scores) contribute substantially to the prediction of academic success (e.g., course grades, honors, acceptance and graduation from graduate or professional school). Additionally, the study found that academic preadmission measures contribute moderately to predicting outcomes with both academic and nonacademic components (e.g., persistence and graduation) and make a small, but significant contribution to predicting college leadership, college accomplishments (e.g., artistic, athletic, business), and post-college income. Given these results, the study concluded that academic variables have practical utility in college admissions in light of their validity in predicting outcomes related to the broader objectives of postsecondary education. However, the researchers also acknowledged that “Although a common metric for evaluating college applicants, standardized tests assess a limited range of academic skills and do not effectively measure non-academic skills" (Burton & Ramist, 2001, p. 4). As discussed later in this chapter, several scholars (e.g, Sedlacek, 2004; Tracey & Sedlacek, 1984) have argued that non-academic skills are critical for supporting college student success.

A comprehensive meta-analysis of 3,000 SAT validity studies reported that SAT validity coefficients for predicting first-year GPA ranged from .44 to .62, when corrected for range restriction and criterion unreliability (Hezlett et al., 2001). This meta-analysis also found that the SAT serves as a valid predictor of academic performance later in college (e.g., graduation, cumulative undergraduate grade point average) with multiple correlations ranging from .35 to .45 (Hezlett et al., 2001). Conversely, the findings of other predictive validity studies (e.g., Bowen & Bok, 1998; Bridgeman, McCamley-Jenkins, & Ervin, 2000; Burton & Ramist, 2001; Sedlacek, 2004; Vars & Bowen, 1998; Wilson, 1983) which investigate the correlation between academic preadmission variables and college grade point average are less conclusive. For example, Camara and Kimmel (2005) suggest, "Although there are a variety of ways to measure general cognitive skills or achievement, there is a limit to the incremental validity to be gained by including additional predictors of the same general construct" (p. 69). Sternberg (1997, 2010) indicated that broadening the measured predictors beyond analytical skills offers a greater promise of enhancing prediction than continuing efforts to improve and refine the same cognitive predictors used in college admissions for decades.

Differential prediction of traditional criteria in college admissions. Given the role of college admissions as a gateway to postsecondary participation, there is applied interest in examining the validity of admission criteria given the extent of variability across groups of college applicants who differ on important demographic variables such as race, ethnicity, gender, and socioeconomic status (Young & Kobrin, 2001). The examination of differential prediction is of particular importance when criteria are used for admission selection decisions (Maller & Pei, 2017). According to Young and Kobrin (2001), “For any educational or psychological test, the validity of the instrument for its intended purposes should be the primary consideration for users of that test” (p. 1). Thus, the reliability of a measure is a prerequisite for its predictive validity.

In the context of college admissions, differential prediction exists when the prediction of criteria is different across groups such that a specific subgroup(s) “benefits” from over-prediction (Aguinis, Culpepper, & Pierce, 2016). Differential prediction and predictive bias refer to differences in the patterns of associations between test scores and other variables for different groups. In other words, predictive bias refers to the systematic under- or over-prediction of criterion performance for test takers belonging to groups differentiated by characteristics not relevant to the criterion performance (AERA, APA, NCME, 2014; Drasgow & Kang, 1984).

The *Standards for Educational and Psychological Testing* (2014) note that test users should select measures that are fair for all test takers. “A test that is fair within the meaning of the Standards reflects the same constructs(s) for all test takers, and scores from it have the same meaning for all individuals in the intended population” (AERA, APA, NCME, 2014, p. 50).

An underlying objective of college admissions is the assessment of directly unobservable constructs such as intelligence and motivation. College admission officers typically infer the extent to which applicants possess an unobservable characteristic based on their responses to items such as essays that presumably reflect the underlying construct (Pendergast, von der Embse, Kilgus, & Eklund, 2017). Thus, it is of critical importance that the constructs assessed in college admissions have the same meaning for all groups. This ensures that mean differences in scores are reflective of true differences on a particular construct and are not the function of statistical differences in measurement properties that are unrelated to the intended construct(s)

(Pendergast et al., 2017). Evidence of construct underrepresentation and construct-irrelevant difficulty (Messick, 1989) is critical for evaluating the validity of a test because groups may differ significantly in terms of their attitudes, experiences, and capabilities that are not explicitly the intended focus of measurement (Linn, 2005).

The extent to which traditional admission criteria depend on an applicant's ethnicity, race, gender, and socioeconomic status in predicting first-year college grade point average has informed the process of identifying more reliable predictors of college student success. Sackett (2005) suggests that "replacing a current measure (e.g., SAT I) with another measure squarely in the cognitive domain (e.g., achievement: SAT II) can be expected to produce comparable results in terms of subgroup differences" (p. 114). In other words, if measures in the cognitive domain yield a predictable and persistent pattern of subgroup mean differences, then the type of measures that could conceivably serve as a replacement must be from a different domain and be minimally correlated with traditional measures (Sackett, 2005).

The potential existence of differential prediction by gender and ethnicity has been examined for several decades. According to Aguinis, Culpepper, and Pierce (2016), in the majority of these studies, differential prediction has been found, on average, to be small such that standardized tests overpredict grades for African American and Hispanic students (e.g., Mattern & Patterson, 2013) and underpredict grades for female students (e.g., Ancis & Sedlacek, 1997). The majority of this body of work has focused on understanding the degree of differential prediction in specific institutions or the average degree of differential prediction across institutions.

A recent study (Aguinis et al., 2016) found that the use of SAT and high school grade point average results in minimal differential prediction only when examining data across samples collected from 176 college and universities. Findings revealed substantial variability in differential prediction at individual institutions within the sample (Aguinis et al., 2016). Similarly, Rothstein (2004) found that the contribution of SAT scores to the prediction of academic performance was reduced when a range of demographic variables, including race and demographic characteristics of the high school a student attended, were statistically controlled. These conclusions serve to substantiate issues related to the differential prediction of traditional admissions criteria. Despite these understandings, their research does not offer a more reliable alternative to the use of SAT scores and high school grade point average for predicting college student success.

Young and Kobrin (2001) conducted a meta-analysis of 49 studies that examined differences in prediction for different racial, ethnic, and gender groups. The typical design of these studies used first-year grade point average as the criterion and standardized test scores (most frequently SAT) and high school grades as the predictor variables. Past studies consistently observed small-to-moderate differences in the magnitude of the validity coefficients and in the accuracy of prediction equations across racial, ethnic, and gender groups (e.g., Kobrin, Sathy, & Shaw, 2007; Young & Kobrin, 2001). For example, Kobrin, Sathy, and Shaw (2007) found that subgroup differences in SAT, ACT, and the National Assessment of Educational Progress (NAEP) are consistently greater as compared to high school grade point average. However, it should be noted that high school grade point average has the greatest amount of inconsistency (i.e., the lowest reliability) given the variability of its measurement.

In general, previous research has found smaller correlations between SAT scores and first-year college grade point average for African American and Hispanic students compared to White students (Kobrin, Patterson, Shaw, Mattern, & Barbuti, 2008; Young & Kobrin, 2001).

Conversely, the correlation between SAT scores and first-year college grade point average is generally slightly higher for female students than male students (Kobrin et al., 2008). In sum, the literature indicates that SAT scores underpredict college success for female students and overpredict college success for underrepresented students, with the exception of Asian students, where the results provide minimal explanation (Kobrin et al., 2008).

In addition to differential prediction across racial, ethnic, and gender groups, implicit in the criticism of standardized tests and their use in high-stakes decision-making (e.g., college admissions) is that socioeconomic status has an artificial and irrelevant effect on test scores (Sackett et al., 2009). This criticism suggests that through the knowledge of test-taking techniques, for example, college admission applicants of higher socioeconomic status earn higher test scores, but do not necessarily have greater aptitude on the characteristic the test intends to measure (e.g., developed abilities relevant to academic performance). This assertion is comparable to claims that socioeconomic status has a similar effect on academic performance measures in which grading favors students of higher socioeconomic status. This results in artificial test-grade correlations and the evidence of a valid measure. Socioeconomic status correlates to the development of abilities that are predictive of academic performance, although it is not the case that test validity is an artifact of socioeconomic status (Sackett et al., 2012).

Beyond an acceptable level of predictive validity, statistical differences and educational policy considerations inform the types of assessments used in college admissions (Atkinson & Geiser, 2009). For this reason, scholars have argued that standardized admissions tests should be criterion-referenced rather than norm-referenced. In other words, standardized tests would have greater validity for use in college admissions if scores demonstrate applicants' proficient mastery of college preparatory subjects to benefit from and succeed in college as opposed to its current use as a criterion for comparing admission applicants (Atkinson & Geiser, 2009).

Limitations of traditional criteria in college admissions. A separate, but related body of literature revealed limitations of the traditional criteria used in college admission decision-making. These limitations include the variability of high school grading standards and course rigor (Atkinson & Geiser, 2009; Buckley et al., 2018; Burton & Ramist, 2001; Camara & Kimmel, 2005; Syverson, 2007; Zwick, 2002) as well as disparities in postsecondary preparation (Zwick, 2002).

Proponents of the use of standardized tests in college admissions cite the need to control for the variability of high school grading standards as well as the availability and quality of postsecondary preparation (Atkinson & Geiser, 2009; Buckley et al., 2018; Burton & Ramist, 2001; Syverson, 2007; Zwick, 2002). Previous research has attributed standardized test performance to the quality of K-12 education across districts and states, as evidenced by teacher-student ratios, spending per student, teachers' education level and years of experience, and school resources on standardized test performance (Zwick, 2002). For example, Syverson (2007) suggests:

> …because students from more affluent and more sophisticated backgrounds are disproportionately likely to use test-preparation resources and pay for taking the test multiple times (thereby likely improving their scores), test results are less likely to constitute a level-playing-field standard by which to reduce the impact of differences in high school grading standards and academic rigor, thereby reducing the original positive impact of the tests. (p. 60)

To examine the efficacy of the criteria traditionally used in college admissions, one must consider the origins and intent of the current body of predictive validity literature. Despite the substantial depth of research, published studies on the predictive validity of standardized tests are predominantly limited to organization-sponsored (e.g., College Board, Educational Testing

Service, NACAC) studies that argue for or against the utilization of standardized tests in college admissions. For example, the research of Hiss and Franks (2014) was endorsed by NACAC while the work of Atkinson and Geiser (2012), Kobrin and Patterson (2011), Burton and Ramist (2001), Morgan (1989), and Sackett et al. (2009) was commissioned by The College Board, a leading standardized test publisher. The literature also includes institution-specific studies that aim to inform future admissions practices and policies.

Another limitation of previous predictive validity studies is the narrow scope of the institutions examined. Selective liberal arts institutions have conducted the majority of institutional predictive validity studies. As such, the research on the predictive validity of cognitive assessment in college admissions lacks external validation. While broader in the range of institutions considered, the research of Hiss and Franks (2014) includes six public institutions that offer "automatic" or "guaranteed" admission to all students who meet the criteria of high school performance, regardless of the results of their standardized test scores. As such, further research is required to determine the predictive value of traditional admission criteria at public institutions of higher education, especially those with test-optional admission policies.

Efficacy of Test-Optional Admissions Policies

A growing body of research has examined the efficacy of test-optional admission policies in advancing the goals of higher education by expanding access to socioeconomically disadvantaged, underrepresented, and first-generation college students (e.g., Belasco et al., 2015; Hiss & Franks, 2014; Mulugetta, 2013, Shanley, 2007; Soares, 2012). Belasco, Rosinger, and Hearn (2015) conducted a longitudinal study to investigate the impact of implementing a test-optional policy on low-income and underrepresented students in selective liberal arts colleges. Their findings indicated that test-optional admissions policies are limited in addressing manifest goals of expanding educational opportunity for low-income and underrepresented students. Belasco et al. offered evidence that suggests test-optional admissions policies fulfill a latent function of increasing the perceived selectivity and status of selective liberal arts institutions. Test-optional admissions policies typically increase the overall applicant pool, allowing institutions to deny more applications and thereby appear more selective. In doing so, institutions can better control the average standardized test scores of enrollees, a criterion often considered as part of the methodology used to rate institutions such as the rankings published by *U.S. News & World Report*. Therefore, test-optional policies may serve to reproduce and maintain the current social structure and its inequalities within American higher education (Belasco et al., 2014).

Counter to the findings of Belasco et al., a study of the outcomes of a test-optional admissions policy at Providence College identified a 31.1% increase in enrollment of ethnic minority students and 18.8% in first-generation college students (Shanley, 2007). Similarly, a study of test-optional admissions at Ithaca College revealed that the implementation of a test-optional admission policy resulted in a 22% increase in freshmen student enrollees from ethnic minority groups (Mulugetta, 2013). Mulugetta (2013) also found that standardized tests contributed 1.8% in predictive validity. Bates College, widely recognized as the first test-optional institution in the United States, introduced a test-optional admission policy, which resulted in a more diverse student body, with an increase in applications and enrollment of ethnic minority students of more than 100% in the first five years (Clinedinst & Hawkins, 2009). These findings suggest that students of ethnic minority groups pursue test-optional admission policies at a higher rate than the total applicant pool (Soares, 2012). However, these colleges reflect a limited subset of institution-types and are therefore not representative of the current landscape of American higher education.

Predictive Validity of Non-Cognitive Variables in College Admissions

In order to attain the broader goals of higher education, researchers have emphasized the need to develop valid and unbiased predictors of college student success (Sackett et al., 2001; Thomas, et al., 2007). To this end, previous research has attempted to identify and validate variables with minimal correlation to the traditional criteria used in college admissions in order to account for the unexplained variance in college performance across student populations. These studies have focused on the use of selection criteria that assess a more complete range of relevant attributes of college applicants (e.g., Komarraju et al., 2009; Thomas, et al., 2007; Tracey & Sedlacek, 1985). Burton and Ramist (2001) posit, “Supplementing the current academic measures of success in college with some of the more important nonacademic measures of success and adding the appropriate nonacademic predictors— such as motivation, interests, or extracurricular accomplishments—would further improve the validity of admission decisions” (p. 5). Furthermore, Camara and Kimmel (2005) suggest, “Research on noncognitive measures appears to be the most promising avenue to pursue at this time if one is interested in incremental validity, constructs that are most distinct from widely used cognitive measures, and different distributions of test takers on a wider range of abilities” (p. 82).

Non-cognitive measures encompass a wide variety of constructs. The following are general domains of non-cognitive variables: applicant characteristics; personal qualities and experiences; and temperament and personality traits (Camara & Kimmel, 2005). Applicant characteristics are variables associated with a student’s background such as state of residency, high school attended, gender, race and/or ethnicity, and socioeconomic status (Camara &

Kimmel, 2005; NACAC, 2016). Personal experiences and biographical data provide important information about future performance and accomplishments, as typically there is consistency in an individual's biological and demographic characteristics. As discussed in detail later in this chapter, temperament and personality traits refer to the distinguishing qualities or characteristics of individuals as evidenced by their habitual patterns of behavior, temperament, and emotion. The conceptual framework of the current study draws from several domains of non-cognitive factors including temperament, personality traits, and self-appraisals.

Applicant characteristics and biographical data. Colleges and universities typically consider a variety of unchangeable personal characteristics and biographical information when evaluating and selecting admission applicants (Linn, 2005). The variables often include, but are not limited to, race, ethnicity, age, gender, and socioeconomic status (as commonly approximated by zip code and/or Expected Family Contribution [EFC]). Studies have found that these non-cognitive variables correlate with first-year college grade point average (e.g., Sackett et al., 2012; NACAC, 2016; Walpole, 2003; Zwick, 2002). Additionally, postsecondary-situational factors such as the level of financial aid received, employment status, family obligations, and out-of-state enrollment may be predictive of college student success and persistence (Allen et al., 2010). According to Lucido (2014), it is normative practice that family financial resources are considered in admission decision-making. Therefore, "need-aware" institutions may use EFC or similar financial indicators as admission selection criteria (Brooks, 2014).

Although decades of legal cases (e.g., *Fisher v. University of Texas at Austin*, 2013; *Gratz v. Bollinger*, 2003; *Grutter v. Bollinger,* 2003; *University of California Regents v. Bakke*, 1978) have considered the extent to which institutions can use biographical information for selecting admissions applicants, many institutions continue to place value on these factors for shaping incoming student classes. Certainly, there is a compelling case for the benefits of diversity to institutions, broader society, and to both ethnic majority and minority students (Bowen & Bok, 1998). These benefits do not exclusively derive from greater aggregate academic performance, but from experiences that are more robust for all students (Linn, 2005). As such, applicant characteristics will continue to inform admission decision-making methodologies unless prohibited by future legislation action.

Personal qualities and experiences. Comprehensive and holistic evaluative methodologies in college admissions traditionally consider contextual factors associated with the size, demographics, academic rigor, and competiveness of the high school from which the applicant graduated (Atkinson & Geisier, 2009; Ting & Robinson, 1998). These factors may lend insight into an applicant's ability to adapt to the academic and social demands of postsecondary education. Sternberg (2010) studied the use of non-cognitive variables to measure an individual's ability to interpret information and adapt to changing environments. This ability influences how college students transition to different learning environments and learn to manage the circumstances they experience (Sedlacek, 2004). Similarly, the personal qualities and experiences of college applicants lend insight into their ability to benefit from college, potential to contribute to society, and/or capacity to overcome adversity (Bowen & Bok, 1998; Camara & Kimmel, 2005; Perfetto, Escandón, Graff, Rigol, & Schmidt, 1999).

Scholars who have advanced broader definitions of college success point to other personal qualities of college applicants that support their academic achievement (e.g., Campbell, McCloy, Oppler, & Sager, 1993; Willingham, 1985). These factors include personal growth, intellectual growth, ability to communicate, and the ability to master new concepts (Willingham, 1985). Similarly, Campbell, McCloy, Oppler, and Sager (1993) identified three primary dimensions of proficiency that apply to educational achievement: declarative knowledge (e.g., facts, principles, and self-knowledge), procedural knowledge and skill (e.g., knowledge and skill of how to accomplish the work, cognitive skill, interpersonal skill, etc.), and motivation (e.g., level of effort, persistence). Applicants acquire these skills through personal experiences and participation in formal educational activities such as classroom-based learning.

Unlike unchangeable biological characteristics, personal qualities and experiences are more easily influenced by interventions such as the development of study skills (Nisbet, Ruble, & Shurr, 1982), extracurricular involvement (Astin, 1993), and academic and social integration (Milem & Berger, 1997; Tinto, 1993). These interventions are associated with fundamental components of student retention models (Tinto, 1993) and are therefore of interest in predicting student persistence in college. Specifically, Tinto (1975) posited that levels of academic and social integration are strong predictors of student persistence in college. Despite the extensive range of non-cognitive variables, the current body of literature provides few consistent and definitive understandings of the relationship between lived experiences and academic achievement in college.

Personality traits, temperament, and self-appraisal. A central objective of college admissions is the prediction of future performance. Personality scales, often implemented as self-report instruments, typically intend to predict future behaviors as evidenced by current behaviors, temperament, emotional patterns, attitudes, and values that evolve from the aforementioned biological and environmental factors. Although frequently utilized in educational and psychological settings, institutions rarely use personality scales for the evaluation of admission applicants. The limited use of personality scales in college admissions reflects the need for evidence of the extent to which personality traits relate to academic motivation and achievement (Komarraju et al., 2009). Thus, there is applied interest in validating measures of personality traits and examining their utility for evaluating college admission applicants.

Despite these limited understandings, several scholars have investigated whether personality traits (particularly those described by the Five-Factor Model of Personality) relate to a wide range of behaviors and subsequent outcomes such as academic achievement (Komarraju et al., 2009) and job performance (Costa & McCrae, 1992; Judge, Jackson, Shaw, Scott, & Rich, 2007). Table 2.1 includes brief descriptions of the Big Five Personality Traits.

Table 2.1:

Big Five Personality Traits

Big Five Personality Trait	Description
Neuroticism	Level of stability versus instability
Extraversion	Tendency to be assertive, sociable, energetic, and outgoing
Openness	Disposition to be curious, open to new and unconventional situations, and imaginative
Agreeableness	Disposition to be cooperative, supportive, easy to get along with
Conscientiousness	Disposition toward purposeful, determined, and goal-directed behavior

Note. Source: "Choosing Students: Higher Education Admissions Tools for the 21st Century," by W. J. Camara and E. W. Kimmel, 2005, p. 87, Mahwah, NJ: Lawrence Erlbaum Associates. Copyright 2005 by Lawrence Erlbaum Associates.

Extensive empirical evidence supports the use of the Five-Factor Model of Personality as an effective unified theoretical framework for the study of personality in different settings with diverse populations (Camara & Kimmel, 2005; Costa & McCrae, 1994; Digman, 1997; Komarraju et al., 2009). For example, Noftle and Robbins (2007) found that conscientiousness predicted college grade point average, even after controlling for high school grade point average and SAT scores, and that the relation between conscientiousness and college grade point average was mediated, both concurrently and longitudinally, by increased academic effort and higher levels of perceived academic ability (Noftle & Robbins, 2007). Similarly, Komarraju et al., (2009) identified patterns of relationships between personality traits and academic motivation. These findings suggest that personality traits influence academic motivation and performance, even after controlling for traditional predictors of academic achievement in college. However, further research is required to determine the predictive validity of the scales associated with the Five Factor Model and their use in the context of college admissions. Given its promise in employment and educational settings, there is applied interest in the validity of personality traits in predicting college student outcomes (Camara & Kimmel, 2005; Komarraju et al., 2009; Noftle & Robbins, 2007).

Despite these understandings, research on non-cognitive predictors is limited and primarily involves small-scale studies with samples of students from a single institution (Camara & Kimmel, 2005). According to Camara and Kimmel (2005), "The absence of empirical evidence of the validity and reliability of other non-cognitive measures of college performance remains a major obstacle to both scientific efforts and the academic and economic incentives needed to sustain a long-term commitment to the needed research" (p. 58). As such, investigation of the incremental validity of a broader range of non-cognitive variables in predicting college outcomes is required.

Predictive Validity of Motivational-Developmental Dimensions in College Admissions

A subset of the research on non-cognitive constructs has investigated a range of psychosocial factors including emotional intelligence, personality traits, coping skills, and achievement motivation. Given the prominence of motivation as a dimension of proficiency related to educational achievement, there is applied interest in evaluating the validity of motivational constructs as predictors of college student success, particularly academic performance and student retention.

Several scholars have cited the promise of motivational constructs as predictors of college outcomes. For example, a large body of research indicates that self-efficacy correlates with academic performance (e.g., Bandura, 1986; Schunk, 1981; Zimmerman, 1981). Additionally, Robbins et al. (2004) suggest, "The evidence is strong that motivational constructs are associated with college performance" (p. 264). Furthermore, Zwick (2002) suggests:

> Attaining campus diversity in the immediate future requires flexible admissions policies that allow for the consideration of a wide range of student characteristics. Among the student attributes that warrant further investigation are motivation, perseverance, and 'spike talents' in particular areas. Such factors are already considered at many schools, but we need more research to determine how best to measure these characteristics and to assess their predictive value. (p. 141)

Thus, a broader range of admission measures may also improve prediction for those students who exhibit motivational characteristics and use coping mechanisms to succeed in college (Burton & Ramist, 2001). Therefore, there is need to investigate motivational constructs and their ability to capture the variety of student characteristics related to success in college. Moreover, the use of non-cognitive factors such as motivational constructs may expand the predictive potential derived from the most frequently utilized admission variables such as high school grade point average and standardized tests (Sedlacek, 2004).

While many factors that influence postsecondary success cannot easily be changed, academic motivation and academic self-concept influence predictors of academic success (Pintrich & DeGroot, 1990; Robbins, Allen, Casillas, Peterson, & Le, 2006). Several scholars have consistently argued that intelligence is a constant, unchangeable construct and does not necessarily reflect an individual's potential for future growth (e.g., DiPerna & Elliott, 1999; Gardner, 2011; Sternberg, 1985). Thus, there are competing theories of intelligence that portray a personal attribute such as intelligence as relatively fixed (e.g., entity theory) or relatively malleable (e.g., incremental theory; Dweck, Chiu, & Hong, 1995; Dweck & Legget, 1988). The purpose of the current study is not to examine the efficacy of these theories, but rather examine the extent to which motivation predicts future performance, which as scholars have noted, is also a product of personal attributes such as intelligence.

A meta-analysis conducted by Robbins et al. (2004) identified and examined the following constructs derived from educational persistence and motivational models of college success: achievement motivation, academic goals, institutional commitment, perceived social support, social involvement, academic self-efficacy, general self-concept, academic-related skills, and contextual influences. After controlling for the effects of traditional admissions variables and socioeconomic status, the researchers determined that the following psychosocial constructs demonstrate incremental validity in predicting academic performance (i.e., cumulative college grade point average) and persistence (i.e., college retention): academic self-efficacy, achievement motivation, and academic goals. The researchers found six additional constructs to be predictive of the persistence criterion beyond the effects of the traditional predictors:

academic goals, academic self-efficacy, institutional commitment, academic-related skills, social support, and social involvement (Robbins et al., 2004).

In an earlier study, Willingham (1985) described the motivational constructs that contribute to the prediction of academic success beyond the variance explained by the exclusive use of academic measures. For example, he found that the construct “follow through” (defined as a continuous commitment to success in one or more areas) had a particularly strong correlation to subsequent academic performance (Perfetto et al., 1999). According to Willingham (1985), students who have insight into themselves and are able to communicate this to others in a creative way tend to do better in college. His finding suggests that personal statements may serve as a useful tool for collecting data on the emotional intelligence and motivation of college applicants.

Although higher education research has examined social and personal dimensions since the 1950s, several scholars have more recently attempted to develop and validate scales that utilize motivational constructs and personality traits for predicting the academic performance and persistence of college students (Sedlacek, 2004). These scales include, but are not limited to, the Noncognitive Questionnaire (NCQ; Tracey & Sedlacek, 1984) and Student Readiness Inventory (Le, Casillas, Robbins, & Langley, 2005).

Tracey and Sedlacek (1984) expanded upon earlier research on the assessment of social and personal dimensions of college students by identifying eight non-cognitive dimensions used in their development of the NCQ, a scale designed to capture the broader attributes of college applicants, particularly across student populations (Thomas et al., 2007; Tracey & Sedlacek, 1984). Table 2.2 describes the eight dimensions of the NCQ.

Table 2.2:

Non-Cognitive Questionnaire (NCQ) Subscale

NCQ Subscale Dimension	Definition
Positive self-concept	Demonstrates confidence, strength of character, determination, and independence
Realistic self-appraisal	Recognizes and accepts any strengths and deficiencies, especially academic, and works hard at self-development; recognizes need to broaden his or her individuality
Successfully handling the system (racism)	Exhibits a realistic view of the system on the basis of personal experience of racism; committed to improving the existing system; takes an assertive approach to dealing with existing wrongs, but is not hostile to society and is not a "cop out"; able to handle a racist system
Preference for long-term goals	Able to respond to deferred gratification; plans ahead and sets goals
Availability of strong support person	Seeks and takes advantage of a strong support network or has someone to turn to in a crisis or for encouragement
Leadership experience	Demonstrates strong leadership in any area of his or her background
Community involvement	Participates and is involved in his or her community
Knowledge in acquired field	Acquires knowledge in a sustained or culturally related way in any field

Note. Source: "Beyond the Big Test: Noncognitive Assessment in Higher Education," by W. E. Sedlacek, 2004, p. 37, San Francisco: Jossey-Bass. Copyright 2004 by John Wiley & Sons.

Thomas et al. (2007) used meta-analytic methods to examine the predictive validity of scores on the NCQ. The researchers concluded that the NCQ is not a valid predictor of student success in college. Their findings suggest that half of the NCQ subscales failed to make a significant incremental contribution to predictive validity despite large sample sizes and the fact that the sample demonstrated representativeness by consisting of entering freshmen (Thomas et al., 2007).

Despite the findings of Thomas et al. (2007), several constructs measured by the NCQ require further investigation given their prominence in the literature as predictors of college success. These constructs include positive self-concept, realistic self-appraisal, and preference for long-term goals. According to Sedlacek (2004), “A positive self-concept is predictive of success in higher education for students of color and other nontraditional students” (p. 39).

The aforementioned research of Robbins et al. (2004) provided a framework for the development of the Student Readiness Inventory (Le et al., 2005), which aims to identify psychosocial and skill factors that are predictive of college student success and validate the motivational constructs meta-analytically examined by Robbins et al. (2004). Despite affirming the findings of Robbins et al. (2004), Le et al. (2005) acknowledged that “Future study is needed to explore the criterion and predictive validities of the factors constituting this inventory” (p. 482).

Similarly, a meta-analysis conducted by Allen, Robbins, and Sawyer (2010) found that “Measures of psychosocial factors are incrementally predictive of college outcomes and significantly contribute to the proportion of an outcome’s explained variance beyond what is explained by traditional predictors” (p. 1). However, the differential roles of achievement motivation and academic goal constructs are less clear (Allen et al., 2010).

Conceptual Framework

According to Sedlacek (2004), the use of motivational constructs may expand the predictive potential derived from the variables most frequently used in college admissions such as high school grade point average and standardized tests. A central hypothesis of the current study is that scores on a measure of motivational-developmental dimensions including academic motivation, self-concept, academic goal orientations, and academic self-efficacy are valid and reliable predictors of college student success and therefore have utility in college admissions. As such, non-cognitive factors such as emotional intelligence, personality traits, and motivation may have predictive utility for college admission officers to select applicants with demonstrated potential for success in college.

The body of research that addresses the identification of variables used to predict college student success has both practical and theoretical significance. Le et al. (2005) suggest, “Although the educational literature is replete with theories and models proposing various psychosocial and/or skill factors that may influence students’ college success (Pascarella & Terenzini, 1991; Tinto, 1975, 1993), the diversity of the constructs and the lack of an integrative framework limit the development of a multidimensional inventory with strong psychometric and conceptual underpinnings” (p. 483). Thus, the current study aims to contribute to the development of this framework. Table 2.3 includes key motivational constructs as defined by Robbins et al. (2004).

Table 2.3:

Key Motivational Constructs as Defined in the Literature

Motivational Construct	Definition
Achievement Motivation	"One's motivation to achieve success; enjoyment of surmounting obstacles and completing tasks undertaken; the drive to strive for success and excellence."
Academic Goals	"One's persistence with and commitment to action, including general and specific goal-directed behavior, in particular, commitment to attaining the college degree; one's appreciation of the value of college education."
Academic Self-Efficacy	"Self-evaluation of one's ability and/or chances for success in the academic environment."
General Self-Concept	"One's general beliefs and perceptions about him/herself that influence his/her actions and environmental responses.".

Note. Source: "Do Psychosocial and Study Skill Factors Predict College Outcomes?" A Meta-Analysis, by S. Robbins, K. Lauver, H. Le, D. Davis, R. Langley, A. Carlstrom, and H. Cooper, 2004, p. 267. Copyright 2004 by American Psychological Association.

The current study examined six central motivational-developmental dimensions found to relate to college students' persistence, academic performance, and retention: self-concept, self-set goals, causal attributions, coping strategies, self-awareness, and self-authorship. These constructs reflect the need for college students to have motivational resources to facilitate effort and persistence, and developmental maturity to apply these resources adaptively in the context of college (Kaplan, Pendergast, French, & Kanno, 2017). The social-cognitive motivational perspective (Bandura, 1986, 2006) and the constructivist-developmental perspective (Kegan, 1994) provide corresponding complementary theoretical frameworks from which the constructs were drawn.

According to Bandura (2006), efficacy beliefs

> …influence whether people the courses of action people choose to pursue, the challenges and goals they set for themselves and their commitment to them, how much effort they put forth in given endeavors, the outcomes they expect their efforts to produce, how long they persevere in the face of obstacles, their resilience to adversity, the quality of their emotional life and how much stress and depression they experience in coping with taxing environmental demands, and the life choices they make and the accomplishments they realize. (p. 309)

In other words, the social-cognitive perspective underscores the contribution to persistence, adjustment, coping, and performance of the combined influence of students' high competence perceptions; attributions of success and of failure to internal, malleable, and controllable causes; self-setting of autonomous, challenging, specific, and realistic goals; and coping with difficulties and failure by focusing on analyzing the problem, regulating negative emotions, and applying context-specific strategies (Bandura, 2006).

The conceptual framework of the current study includes four motivational constructs based on the social-cognitive perspective: self-concept (i.e., individuals' self-perceptions of ability [Marsh & Martin, 2011]); self-set goals (i.e., an individual's personally determined goals for themselves and for others [Locke & Latham, 2002; Vansteenkiste et al., 2013]); causal attributions (i.e., cognitive-affective explanations of the causes of success and failure [Hong et al., 1999; Weiner, 2010]); and coping strategies (i.e., purposeful behavioral, emotional, and cognitive actions for responding to situations perceived to challenge an individual's resources [Compas, Connor-Smith, Saltzman, Thomsen, & Wadsworth, 2001]). These social-cognitive constructs are related. However, they constitute distinct attributes that combine to form an adaptive motivational mindset (Dweck & Leggett, 1988).

The developmental-constructivist perspective reflects the role of cross-contextual capacities for intentional and purposeful self-reflection and self-regulation of knowledge, relationships, goals, and actions related to coping and growth (Kegan, 1994). The conceptual framework of the current study includes two developmental constructs based on the developmental-constructivist perspective: self-awareness (i.e., the ability to consider oneself as an object for reflection, monitoring, and learning [Silvia & Duval, 2001]) and self-authorship (i.e., the agentic capacity for an individual to generate and regulate their beliefs, decisions, identity, and social relationships [Baxter-Magolda, Creamer, & Meszaros, 2010]).

Informed by the literature as well as theoretical, empirical, ethical, and logistical factors, Kaplan et al. (2017) advanced the following definitions of the selected motivational-developmental dimensions that served as the conceptual framework of the current study.

Attributions. Attributions reflect the characteristics of someone who accredits success and failure to controllable and malleable causes (e.g., effort as well as strategies). Attributions are "cognitive explanations of the causes of events in the world that reflect the person's ontological and epistemological explicit and implicit assumptions. When concerned with the person's own experiences of success and of failure, attributions have been related to emotions, future expectations of success and failure, motivation, and achievement."

Coping. Coping reflects the characteristics of someone who "manifests problem-focused coping that is specific and situation appropriate as well as adaptive recruitment of social supports during experiences or challenge." Coping strategies refer to "active and purposeful processes by which an individual responds to taxing events or situations that are perceived to exceed his or her resources. Coping includes behavioral, emotional, and cognitive attempts to manage the demands imposed by a stressor."

Future time perspective. Future time perspective refers to the "constructing of a view of the self in the future. This construction is characterized by extent (length of time) and content (the nature of the self-image). Generally, people who construct a longer future are more motivated to pursue it and are able to view the value of present efforts towards the future than people who construct a shorter future. People who are more optimistic about themselves in the future are more motivated and adjusted than people who are more pessimistic. People who construct an optimistic and elaborate view of themselves in the future have higher need for achievement, are more persistent in obtaining goals, have stronger ability to delay gratification, and suffer less from emotional distress and hopelessness."

According to Lens (1986), by setting goals in the near or distant future, individuals develop their specific future time perspective that can be characterized by its content (i.e., what individuals strive to achieve) and extension or depth (i.e., how far into the future individuals set their goals). Individual differences in the content and extension of the future time perspective may have motivational consequences and therefore influence student success outcomes.

Relevant experiences. Relevant experiences refer to characteristics of someone who describes pertinent understandings and life events and is able to "explicate realistic and sophisticated connections from these experiences to future imagined college experience…"

Self-authorship. Self-authorship refers to the characteristics of someone who "manifests agency in decision-making that involves attending to norms and others' advice as information and a strong personal voice in self-exploration and decision-making." Self-authorship also refers to a "comprehensive outlook on one's role in life as an 'author'—being active in taking responsibility for choosing beliefs, constructing a sense of self in context, setting a course of action, and forming and influencing social relationships. It manifests in having an internal voice that evaluates and coordinates meaning of experiences about one's life in areas of beliefs, role in different contexts, and social relationships."

Self-awareness. Self-awareness refers to the characteristics of someone who is "self-reflective and aware of the multidimensionality of self and of context." Self-awareness also refers to the "state of considering oneself as an object for reflection, examination, and learning. A developed self-awareness involves elaborate reflection on one's own traits, strengths and weaknesses, thoughts, emotions, behavior, certainties and uncertainties, and changes (e.g., during a given length of time and as influenced by context, situations, and social interactions)."

Self-concept. Self-concept refers to the characteristics of someone who is "optimistic and confident but not unrealistically overconfident or self-aggrandizing." This dimension also refers to someone who "holds a set of adaptive beliefs about success" and recognizes their own limitations. Additionally, self-concept encompasses a "person's self-perceptions of ability in a domain (e.g., academics, social, physical). Self-concept relates to an individual's self-esteem and perceived self-worth. Self-concept also relates to an individual's self-efficacy and perception that he or she has the ability to organize and execute successfully a course of action required to attain a certain level of performance on a particular task."

Self-set goals. Self-set goals refer to the characteristics of someone who is "motivated by specific personal growth goals of learning and developing, intrinsic goals of meaning and interest, extrinsic goals of accomplishment, as well as other-oriented goals." This dimension also refers to the ability to connect past experiences to future imagined college experience. Self-set goals also refer to a "personally determined and desired goal, objective, or purpose for the self

(self-oriented) and for the effect of one's actions on others (other-oriented). Adaptive self-set goals are well-defined, relatively specific, realistic, and are defined in ways that allow assessing progress towards their accomplishment. Self-set goals are meaningful to an individual and involve the desire to make a positive difference in the world."

Given these theoretical definitions, the aforementioned dimensions are reflective of college student motivation (attributions, coping, self-concept, self-set goals) and development (future time perspective, relevant experiences, self-authorship, self-awareness).

CHAPTER 3

METHODOLOGY

The current study investigated the predictive validity of a measure of motivational-developmental dimensions in predicting college student success as measured by cumulative undergraduate grade point average and retention. The aggregate score of eight motivational-developmental dimensions (attributions of successes and failures, coping, future time perspective, relevant experiences, self-authorship, self-awareness, self-concept, and self-set goals) served as the predictor variable (test-optional essay rating score) in the current study. Cumulative undergraduate grade point average and retention served as the outcome variables.

The current study was comprised of two sub-studies (hereinafter referred to as Study 1 and Study 2) in order to account for changes to the methodologies used to evaluate the aforementioned motivational-developmental constructs as part of the admission process at the participating institution. Specifically, Study 1 included the dimension “relevant experiences” while Study 2 replaced this construct with “future time perspective.” Additionally, Study 1 and Study 2 employed different scales used to score the essays and the underlying motivational-developmental dimensions.

Participants

In total, the current study consisted of 10,149 undergraduate students who enrolled as first-time freshmen at a large, urban public university located in the Northeastern United States. Participants of the current study included all enrollees of the Fall 2015 (Study 1) and Fall 2016 (Study 2) entering undergraduate freshmen classes at the participating institution. Study 1 consisted of 4,920 participants including 886 participants who submitted a test-optional admission application. Study 2 consisted of 5,229 participants including 724 participants who submitted a test-optional admission application. Descriptive data on the participants in the two studies are presented in Chapter 4 (Tables 4.1 and 4.2).

Measures

Informed by the literature and theoretical, empirical, ethical, and logistical factors, Kaplan et al. (2017) developed a measurement instrument for the participating institution to assess the preadmission motivation of test-optional admission applicants. The participating institution administered this instrument as part of the test-optional admission process. The instrument consisted of four short-answer essay questions, each of which assessed a subset of the motivational-developmental dimensions (see Tables 3.1 through 3.4). Kaplan et al. (2017) developed comprehensive rubrics to advance a shared definition of the motivational-developmental dimensions as assessed in each of the test-optional essay questions.

Motivational-developmental dimensions. In Study 1, the following motivational-developmental dimensions were assessed in the test-optional essay questions: attributions of success and failure, coping, self-set goals, relevant experiences, self-authorship, self-concept, and self-awareness. Table 3.1 includes descriptions of the essay questions and the corresponding dimensions assessed in each question. Table 3.2 indicates the primary motivational-developmental dimensions assessed in each question. The specific essay questions are not included in the current study given their proprietary nature.

Table 3.1:

Test-Optional Essay Questions and Motivational-Developmental Dimensions in Study 1

Test-Optional Essay Question Description	Motivational-Developmental Dimensions
Essay 1 required applicants to reflect on past success and sources of pride. Applicants discussed how these experiences might relate to future experiences in college.	Self-Concept Self-Set Goals Attribution of Successes Relevant Experiences
Essay 2 required applicants to reflect on experiences of failure. Applicants discussed how the experience might relate to future experiences in college.	Self-Concept Attribution of Failures Coping Relevant Experiences
Essay 3 required applicants to reflect on experiences in which they faced a significant challenge. Applicants discussed how they approached this challenge.	Self-Awareness Self-Set Goals Coping Self-Authorship
Essay 4 required applicants to reflect on alternatives to college participation. Applicants discussed what they might do instead if not admitted to college.	Self-Awareness Self-Set Goals Coping Self-Authorship

Table 3.2:

Primary Motivational-Developmental Dimensions in Study 1

Dimension	Essay 1	Essay 2	Essay 3	Essay 4
Attributions	X	X		
Coping		X	X	X
Relevant Experiences	X	X		
Self-Authorship			X	X
Self-Awareness			X	X
Self-Concept	X	X		
Self-Set Goals	X		X	X

In Study 2, the following motivational-developmental dimensions were assessed in the test-optional essay questions: attributions of success and failure, coping, future time perspective, self-set goals, self-authorship, self-concept, and self-awareness. Table 3.3 includes descriptions of the questions and the corresponding dimensions assessed in each question. Table 3.4 indicates the primary motivational-developmental dimensions assessed in each question. The specific essay questions are not included in the current study given their proprietary nature.

Table 3.3:

Test-Optional Essay Questions and Motivational-Developmental Dimensions in Study 2

Test-Optional Essay Question Description	Motivational-Developmental Dimensions
Essay 1 required applicants to reflect on experiences and causes of success. Applicants discussed sources of pride. Applicants also discussed how experiences of success might relate to future experiences in college.	Self-Concept Self-Set Goals Attribution of Successes Future Time Perspective
Essay 2 required applicants to reflect on experiences and causes of failure. Applicants discussed how the experience of failure might relate to future experiences in college.	Self-Concept Attribution of Failures Coping Future Time Perspective
Essay 3 required applicants to reflect on experiences in which they had to make an important decision. Applicants discussed how they approached the decision and what they considered during the decision-making process.	Self-Awareness Self-Set Goals Coping Self-Authorship
Essay 4 required applicants to reflect on why they decided to attend college. Applicants discussed alternatives to college participation.	Self-Awareness Self-Set Goals Future Time Perspective Self-Authorship

Table 3.4:

Primary Motivational-Developmental Dimensions in Study 2

Dimension	Essay 1	Essay 2	Essay 3	Essay 4
Attributions	X	X		
Coping		X	X	
Future Time Perspective	X	X		X
Self-Authorship			X	X
Self-Awareness			X	X
Self-Concept	X	X		
Self-Set Goals	X		X	X

New Student Questionnaire (NSQ). The New Student Questionnaire (NSQ; see Appendix A) is an 82-item survey administered to all matriculated undergraduate freshmen and transfer students as part of the new student orientation process. In other words, the participating institution does not use the NSQ to evaluate and select admission applicants. Rather, the participating institution uses the NSQ to monitor trends in student characteristics, attitudes, intentions, and aspirations. The institution also uses data obtained from the NSQ to assess factors related to student success, as well as improve student retention, academic programming, and campus climate. The NSQ asks students about the importance of various academic, social, and administrative areas and about their satisfaction or agreement with each survey item. Specifically, items 75-82 measure motivational variables related to academic achievement (see Appendix B). It is important to note that I selected these items for examination given my interpretation of their intent (the measurement of student motivation).

Table 3.5:

New Student Questionnaire (NSQ) Response Rates as a Percentage of Participants

NSQ Item	Study 1 (n = 4,920)	Study 2 (n = 5,229)
Hard Work	96.2	95.3
Difficult Plan	95.6	95.0
Enjoy Study	96.1	95.0
Time Management	96.3	95.2
Confident	95.8	95.0
Give Up	95.7	94.8
Organized	96.0	95.0
Independent	96.1	95.1

Outcome variables. As mentioned, the test-optional essays intended to assess applicants' articulated motivation to succeed in college. As is consistent with the majority of predictive validity studies in college admissions (e.g., Burton & Ramist, 2001; Kobrin & Patterson, 2011; Le et al., 2005), the current study measured cumulative undergraduate grade point average and retention as indicators of college student success. Cumulative grade point average is the number of quality points earned in all courses attempted divided by the total degree-credit hours in all attempted courses. Retention is a dichotomous variable (e.g., retained or not retained) that refers to a student's enrollment status (i.e., registered or not registered) immediately following the final date by which students can withdraw from courses. Upon the start of each semester, enrolled students have two weeks to withdraw from 16-week courses and one week to withdraw from seven-week courses. Study 1 included retention data for the Fall 2015, Spring 2016, Fall 2016, and Spring 2017 semesters. Study 2 included retention data for the Fall 2016 and Spring 2017 semesters.

Procedures

Participants submitted biographical, demographic, and educational data using a secure online application as part of the undergraduate admissions process. All participants submitted either a "standard" or "test-optional" admission application. Standard applicants were required to submit SAT and/or ACT scores as part of their application for admission. Test-optional applicants were required to submit responses to four short-answer essay questions in lieu of SAT or ACT scores. Although not required for admission to the participating institution, 47.5% of the test-optional applicants completed the SAT. These scores were submitted by test-optional admission applicants and obtained by the participating institution from The College Board. All study participants submitted official high school transcripts. Following their admission and subsequent matriculation, all participants submitted responses to the NSQ as part of the new student orientation process.

Trained readers scored each construct within the test-optional essay questions. The aggregate score of the essays served as the predictor variable "test-optional essay rating score." To enhance interrater and intrarater reliability, the development and distribution of a scoring rubric advanced a shared definition of the motivational-developmental dimensions among the readers. Following a series of trainings and test evaluations, the rubric was calibrated using a Kappa statistical analysis to further improve the reliability of the measure.

Procedures for Study 1. Two trained readers independently scored the primary motivational-developmental constructs within each test-optional essay question (see Table 3.2) according to the following scale: 1 point (does not articulate the construct), 4 points (narrowly articulates the construct), 7 points (generally articulates the construct), 10 points (explicitly articulates the construct). Therefore, each essay question received a total score between four and 40 points per reader. A third reader scored an essay response if there was a variance of five points or more between the scores provided by the two initial readers on a given essay question. In other words, an essay was evaluated by a third reader if there was a gap greater than one interval (1 – 7, 1 – 10, 4 – 10). In this circumstance, the score provided by the third reader was accepted and the scores produced by both initial readers were rejected. Given this scoring methodology, a total score from four to 40 was produced by averaging the scores produced for each essay question.

Procedures for Study 2. Two trained readers independently scored the primary motivational-developmental constructs within each test-optional essay question (see Table 3.4) according to the following scale: 0 (does not articulate the construct), 2 (narrowly articulates the construct), 4 (partially articulates the construct), 8 (generally articulates the construct), 10 (explicitly articulates the construct). Each reader had the discretion to award 1–3 bonus points to recognize exceptional articulations of the motivational-developmental dimensions. Therefore, each essay question received a total score between four and 43 points per reader. However, a third reader scored an essay response if there was a variance of five points or more between the scores provided by the two initial readers on a given essay question. In other words, an essay was evaluated by a third reader if there was a gap greater than three intervals (0 – 8, 0 – 10, 2 – 8, 2 – 10, 4 – 10). In this circumstance, the score provided by the third reader was accepted and the scores produced by both initial readers were rejected. Given this scoring methodology, a total score from 0 to 43 was produced by averaging the scores produced for each essay question.

Data Analyses

In response to the primary and secondary research questions posed in the current study, extant secondary data collected through the undergraduate admissions process of the participating institution were analyzed using quantitative methods.

Preliminary analyses. These statistical procedures, namely parametric tests, are based on the assumption that the data follow a normal distribution in order to make accurate and reliable conclusions (Yap & Sim, 2011). To test the distribution of the data, skewness and kurtosis were examined to verify that values for each variable were greater than -1 and less than 1 (Bulmer, 1979). Outliers were screened and values for all predictor and outcome variables were found to be within range of the maximum and minimum limits of the measures from which they were derived. Appendices E and F present additional descriptive statistics on the range, mean, and standard deviation for the predictor and outcome variables. Given the large sample size, both listwise and pairwise procedures did not result in statistically significant differences (Little & Rubin, 2014). Thus, to defer effects on correlations, pairwise deletion was used to handle all cases of missing data in regression analyses.

Descriptive statistics. To contextualize the findings of the current study, chi-squared tests were conducted to provide descriptive statistics on the study participants as well as the predictor and outcome variables. Chi-squared tests were conducted to determine whether there were significant differences between test-optional and standard admission applicants in terms of gender, race, and age. Separate samples T-tests were used to compare the two applicant groups on preadmission academic performance as measured by high school grade point average and SAT scores. Additionally, frequency data on the intended academic discipline of the study participants were compiled.

Research Question 1: Predictive validity of the test-optional essay rating score. To address Research Question 1 ("Is the test-optional essay rating score a significant predictor of college student success, as measured by cumulative undergraduate grade point average and retention?"), a correlation analysis was used. Pearson product-moment correlation coefficients were computed to examine the relationship between the predictor variable (test-optional essay rating score) and outcome variables (cumulative undergraduate grade point average and semesters retained). To provide context as to the statistical significance of the test-optional essay rating score variable, additional Pearson product-moment correlation coefficients were computed to evaluate the relationship between traditional admission variables (high school grade point average, SAT Total, SAT Verbal, SAT Quantitative, and SAT Writing) and the outcome variables (cumulative undergraduate grade point average and semesters retained). This correlation analysis included the predictor variable data for standard applicants as well as test-optional applicants for whom SAT score data were available.

Incremental prediction of test-optional essay rating score. A sub-question of Research Question 1 asked, "Is the test-optional essay rating score incrementally predictive of cumulative undergraduate grade point average and retention given the variance accounted for by traditional admission criteria (high school grade point average and SAT scores)?" The purpose of this question was to investigate the incremental value, if any, of the test-optional essay rating score in predicting cumulative undergraduate grade point average and semesters retained given the proportion of the variance explained by traditional admission criteria (high school grade point average and SAT scores).

Diagnostic accuracy of the test-optional essay rating score. A sub-question of Research Question 1 asked, "At what test-optional essay rating cut score are test-optional admission applicants predicted to persist in college?” The objective of this question was to investigate the diagnostic accuracy of the test-optional essay rating score in predicting first- and second-year retention. Since retention is a dichotomous variable, a Receiver Operating Characteristic (ROC) analysis was conducted to test the criterion validity of the test-optional essay rating score as a predictor variable.

To provide context related to this type of analysis, ROC is a graphical plot of the sensitivity (true positive rate) vs. 1- specificity (false positive rate) for a binary classifier system. According to Youngstom (2013), optimal prediction outcomes for ROC analysis are 100% sensitivity (all true positives are found) and 100% specificity (no false positives are found). However, these prediction outcomes are not practical given the high-stakes nature of college admissions, for both the applicant and institution. According to Yu, DiGangi, Jannasch-Pennell, and Kaprolet (2010), in practice, a satisfactory classification system should depict a ROC curve learning towards the upper left of the graph, which implies maximum prediction of the actual state. McCluckie (2014) suggests that the key to ROC analysis is its use of a discrimination threshold, which is the point at which the test result goes from one outcome to the other (e.g., negative to positive). Table 3.6 shows the contingency table used when conducting ROC analysis.

Table 3.6:

Contingency Table Used in ROC Analysis

Prediction Outcome	Actual State Positive (p)	 Negative (n)	Total
Positive (p')	True Positive (TP)	False Positive (FP)	Positive (P')
Negative (n')	False Negative (FN)	True Negative (TN)	Negative (N')
Total	Positive (P)	Negative (N)	

Note. Source: "Identifying Students 'At Risk' of Withdrawal Using ROC Analysis of Attendance Data," by B. McCluckie, 2014, *Journal of Further and Higher Education*, *38*, p. 527. Copyright 2012 by Taylor and Francis.

According to Youngstrom (2013), "The area under the curve (AUC) from a ROC analysis quantifies the amount of separation between the distribution score for the two groups of interest, those with and without the criterion diagnosis" (p. 208). The choice of an optimal AUC cut score threshold depends on several factors including the desired admission acceptance rate and the extent to which an institution considers admission applicants' probability of retention. Institutions that intend to maintain a low rate of admission acceptance may place a premium on specificity as a diagnostic confirmation. Conversely, tuition-dependent institutions may desire a predictor that has high sensitivity, given the cost of denying admission to applicants who may retain if given the opportunity to enroll. Thus, the optimal cutoff value is one that minimizes cost.

The participating institution admitted 53.5% of freshmen applicants in Study 1 and 50.1% in Study 2. Given that the average acceptance rate was greater than 50%, the determination of a cut score threshold was based on the assumption that the test-optional essay rating score is both used as a screener with high sensitivity and as a diagnostic confirmation (Youngstrom, 2013).

This way, the institution admits applicants who are correctly predicted to retain in college and conversely denies admission to applicants who are correctly predicted to not retain in college. Thus, the optimal cutoff score at the participating institution maximizes sensitivity and specificity.

Contribution of specific motivational-developmental dimensions to the prediction of student success. An additional sub-question of Research Question 1 asked, "Which motivational-developmental dimensions, if any, contribute most to the prediction of cumulative undergraduate grade point average and retention?" Pearson product-moment correlation coefficients were computed to address this question.

Research Question 2: Differential prediction of the test-optional essay rating score. To address Research Question 2 ("Does the test-optional essay rating score differentially predict cumulative undergraduate grade point average and retention across ethnic and gender subgroups?"), a multiple regression analysis was conducted. This analysis investigated whether the test-optional essay rating score differentially predicted cumulative undergraduate grade point average and/or retention across gender and racial subgroups.

Secondary research questions. Secondary research questions were investigated to provide additional context related to the selected measure of motivational-developmental dimensions and to explore the predictive potential derived from an additional measure of college student motivation, the NSQ. A secondary research question asked, "On average, do test-optional applicants attain a greater/lesser cumulative undergraduate grade point average than what is predicted by their standardized test scores?" To answer this question, cumulative undergraduate grade point average and semesters retained were first predicted using the regression equation produced to answer Research Question 2. These predicted scores were then compared to the actual scores through a paired samples, repeated measures t-test.

An additional secondary research question asked, “Are the New Student Questionnaire (NSQ) items that measure student motivation significant predictors of college student success, as defined by cumulative undergraduate grade point average and retention?” The purpose of this question was to investigate whether NSQ items 75-82 (see Appendix B) are significant predictors of college student success, as defined by cumulative undergraduate grade point average and retention. To answer this question, Pearson product-moment correlation coefficients were computed to evaluate the relationship between the select NSQ items and the outcome variables (cumulative undergraduate grade point average and semesters retained).

To evaluate the amount of agreement between test-optional essay readers, interrater reliability was assessed by computing a Kappa statistic. This analysis investigated the consistency of measurement between readers on the specific motivational-developmental dimensions assessed in each test-optional essay question.

Scope of the Study

The purpose of the current study was to investigate the relationship between preadmission motivational-developmental dimensions and subsequent undergraduate student success in college. Given this objective, the current study included several delimitations including a restriction in range of the degree level and type of participants, classification and geography of the participating institution, and scope of the selected motivational-developmental dimensions. Limitations of the current study include changes to the selected motivational-developmental constructs between Study 1 and Study 2, the narrow definition and measurement of college student success, and the timeframe during which the study was conducted.

Participants of the current study included undergraduate freshmen enrollees, as is consistent with the majority of predictive validity literature. The current study restricted transfer students, graduate students, and denied admission applicants from participation. Therefore, findings do not generalize to these student populations, despite the need for clear and consistent evidence across student levels and types. Additionally, the participants of the current study are not representative of the gender, age, races, and ethnicities of postsecondary participants nationally, which limits the generalizability of findings. Appendix C is a summary of the demographic representativeness of the study participants compared to national postsecondary participants.

The current study included a single institution located in a specific geographic region of the United States. The participating institution reflects a particular institutional classification (i.e., public, urban, comprehensive research university) and therefore is not representative of the range of institutional types that collectively characterize the modern landscape of American higher education. Given this delimitation, the current study lacks external validity, which limits the generalizability of findings.

Despite broader definitions of college student success as defined in the literature (Burton & Ramist, 2001; Camara & Kimmel, 2005; Willingham, 1985), the current study measured success over two academic years for Study 1 and one academic year for Study 2 as measured by cumulative undergraduate grade point average and retention. However, scholars have argued the need for longitudinal assessment of student success outcomes such as post-college employment and earnings, graduate school acceptance, and civic participation (Burton & Ramist, 2001; Willingham, 1985).

There are statistical limitations when considering cumulative grade point average as an outcome variable. Grade predictions beyond the first year present statistical issues because attrition of students at various points during their college career. Cumulative college grade point average is a function of academic success over time and this length of time is variable, but nonrandom given student attrition (Sedlacek, 1986). Therefore, credit hours completed was excluded as an outcome variable given that this variable is a function of the total number of semesters in which participants remained enrolled in college.

Informed by the literature as well as theoretical, empirical, ethical, and logistical factors, the currently study investigated the predictive validity of eight motivational-developmental constructs selected by Kaplan et al. (2017) for the development of test-optional essay questions. Despite the range of constructs examined in the current study, the selected dimensions reflect a subset of the possible variables that may correlate to college student success. For example, the selected dimensions compliment, but do not draw directly from, *achievement-as-goals* (Covington, 2000; Le at al., 2005) motivational theories such as expectancy-value theory (Wigfield & Eccles, 2000). Furthermore, *achievement-as-drive* motivational theories such as academic self-efficacy (Bandura, 1986) and self-determination theory (Ryan & Deci, 2000) are not included among the selected dimensions in the current study.

The current study was comprised of two sub-studies in order to account for changes to the motivational-developmental dimensions. Specifically, Study 1 included the construct "relevant experiences" while Study 2 replaced this construct with "future time perspective." These changes limit the assessment of these dimensions over the duration of the current study.

Trained readers scored the test-optional essay questions. Therefore, an assumed level of human subjectivity influenced these scores. The level of consistency of scoring between and within readers also influence the predictive accuracy of the test-optional essay rating score variable. Additionally, changes to the essay scoring methodology between Study 1 and Study 2 also limit the consistency of findings. Specifically, Study 2 introduced an additional interval score ("partially articulates the construct") to the test-optional essay rating scale. This added a level of specificity to the scoring of the test-optional essays that was not available to the readers in Study 1.

CHAPTER 4

RESULTS

This chapter presents the results of the quantitative analyses described in Chapter 3. This chapter is comprised of seven sections, the first of which is a discussion of assumptions and considerations such as the statistical treatment of missing cases. The second section presents descriptive statistics on the study participants including their demographic characteristics and preadmission academic performance. This chapter also includes sections that correspond to each primary and secondary research question. Lastly, this chapter concludes with the results of an interrater reliability analysis conducted to evaluate the level of agreement between the readers of the test-optional essays.

Considerations of the Dataset

The data analyzed in the current study included demographic, biographic, financial, and academic information collected as part of the undergraduate admissions process at the participating institution. Additionally, participants submitted responses to the New Student Questionnaire (NSQ) upon their matriculation. Both of these collection methods involve participant submitted/reported data. This limits the ability to verify the authenticity of these data.

Despite the source of the academic performance data, several methodological decisions were made to ensure the integrity of the data analyses and dataset. Participants with a self-reported race of American Indian, Multiple Ethnicities, Pacific Islander, or Unknown were categorized as "Other" due to the limited racial representativeness of participants (see Appendix C). Additionally, unparallel designations of race (e.g., African American and White) were retained as these categories reflect those available in the institutional dataset.

Eight participants did not enroll in coursework despite initially confirming their matriculation at the participating institution. Therefore, cumulative undergraduate grade point average and retention data were unavailable for these participants. These cases were handled using pairwise deletion to impute the missing data via an algorithmic method (Little & Rubin, 2014). Furthermore, 48 participants had high school grade point averages that exceeded 4.0. These grade point averages were converted onto a 4.0 scale in order to standardize this variable, as many high schools use grading scales that do not exceed 4.0 (NACAC, 2016; Shanley, 2007).

Test-optional essay rating score data were missing for one participant in Study 1 (0.11%) and four in Study 2 (0.55%). Pairwise deletion was used to handle these cases as the effect did not result in statistically significant differences (Little & Rubin, 2014).

There was a fundamental change to the evaluative methodology in which readers in Study 2 had the option to award bonus points to the total test-optional essay rating score. Readers did not have this option in Study 1. Readers awarded bonus points to 184 participants of Study 2. However, bonus points were excluded from the calculation of the test-optional essay rating score variable. In doing so, a consistent maximum value of 40 points for this variable across the sub-studies was maintained. Furthermore, this decision attempted to reduce the level of subjectivity inherent in the allocation of bonus points. This was a particular concern given that the scoring rubric used in Study 2 did not specify the methodology and criteria for awarding bonus points.

Descriptive Statistics

To contextualize the findings presented in this chapter, Tables 4.1 and 4.2 include descriptive statistics on the demographic characteristics of the study participants. Study 1 consisted of 4,920 participants including 886 participants who submitted a test-optional admission application. Participants were 54.7% female including 22.6% of whom submitted a test-optional admission application. Self-reported ethnicities of all participants were 12.4% African American, 11.2% Asian, 6.7% Hispanic, 54.6% White, and 15.0% "Other."

Table 4.1:

Demographic Characteristics of Study 1 Participants

	Test-Optional Applicants N = 886		Standard Applicants N = 4,034	
Demographic Variable	Frequency	Percentage	Frequency	Percentage
Gender				
Female	607	68.5	2,084	51.7
Male	279	31.5	1,950	48.3
Race				
African American	210	23.7	399	9.9
Asian	97	10.9	456	11.3
Hispanic	89	10.0	243	6.0
White	405	45.7	2,281	56.5
Other	85	9.7	655	16.3
Age				
Less than 17	15	1.6	94	2.4
17 – 18	853	96.3	3,789	93.9
19 – 21	15	1.8	127	3.3
22+	3	.3	15	.04

Study 2 consisted of 5,229 participants including 724 participants who submitted a test-optional admission application. Participants were 55.3% female including 17.6% of whom submitted a test-optional admission application. The self-reported races of all participants were 10.5% African American, 12.0% Asian, 5.7% Hispanic, 59.9% White, and 11.9% “Other.”

Table 4.2:

Demographic Characteristics of Study 2 Participants

	Test-Optional Applicants N = 724		Standard Applicants N = 4,505	
Demographic Variable	Frequency	Percentage	Frequency	Percentage
Gender				
Female	508	70.2	2,384	52.9
Male	216	29.8	2,121	47.1
Race				
African American	149	20.6	399	8.9
Asian	93	12.8	536	11.9
Hispanic	69	9.5	228	5.1
White	352	48.6	2,780	61.7
Other	61	8.5	562	12.4
Age				
Less than 17	13	1.9	82	1.8
17 – 18	694	95.9	4,302	95.5
19 – 21	12	1.7	109	2.4
22+	4	.5	12	.3

Chi-square tests were conducted for each of the demographic variables presented in Tables 4.1 and 4.2. As shown in Tables 4.3 and 4.4, there were significant differences between test-optional and standard admission applicants in terms of gender and race. In Study 1, there was a moderate effect size for gender ($\phi = .13$) and a strong effect size for race ($\phi = .18$), given the interpretations of Cramer's V as advanced by Liebetrau (1983)[2]. Additionally, there was a weak relationship between applicant type and age ($\phi = .06$). Similarly, in Study 2, there was a moderate effect size for gender ($\phi = .12$), a strong effect size for race ($\phi = .16$), and a weak effect size for age ($\phi = .06$). Given these interpretations, test-optional applicants were significantly more female and significantly less White than standard admission applicants. There was a negligible relationship between applicant age and the type of admission application submitted. Tables 4.3 and 4.4 present these findings.

Table 4.3:

Demographic Differences by Applicant Type in Study 1

Demographic Variable	Pearson Chi-Square	Significance	Cramer's V
Gender	83.23	.000	.13
Race	166.37	.000	.18
Age	18.76	.13	.06

[2] Accepted interpretations of Cramer's V indicate that variables with values between .01 and .05 have no or negligible relationship. Values between 0.6 and .10 indicate that variables have a weak relationship. Values between .11 and .15 indicate that variables have a moderate relationship. Values between .16 and .25 indicate that variables have a strong relationship. Values .26 or higher indicate that variables have a very strong relationship.

Table 4.4:

Demographic Differences by Applicant Type in Study 2

Demographic Variable	Pearson Chi-Square	Significance	Cramer's V
Gender	75.06	.00	.12
Race	130.63	.00	.16
Age	17.31	.14	.06

As is consistent with the literature (e.g., Hiss & Franks, 2014; Willingham, 1985), the applicant types (i.e., standard, test-optional) of the study participants varied significantly by intended academic discipline. Appendix D presents the intended academic disciplines as indicated on the admission applications submitted by study participants. As shown, participants with an intended academic discipline of Engineering, Business, or Science were less likely to apply for test-optional admissions as compared to participants with an intended academic discipline of Education, Public Health, or Communications. This suggests that applicants with academic interests in quantitative disciplines are more likely to submit standardized test scores as part of their application for admission. In addition, the findings presented in Table 4.5 and 4.6 support the notion advanced by Willingham (1985) that students are more likely to pursue academic disciplines that best align with their level of academic preparation.

Tables 4.5 and 4.6 present descriptive statistics related to the preadmission academic performance of the study participants (i.e., test-optional and standard admission applicants) as measured by high school grade point average and SAT scores. SAT Total scores were available for 50% of the test-optional admission applicants in Study 1 and 44% in Study 2. SAT Verbal, SAT Quantitative, and SAT Writing scores were available for 45% of the test-optional admission applicants in Study 1 and 38% in Study 2.

As indicated in Table 4.5, test-optional admission applicants have lower means on all variables compared to standard applicants. This would be expected for SAT scores. The difference in high school grade point average is also significant with a small effect size (F = 80.89, p = .00, ηp^2 = .016). Consistent with these data for Study 1, the test-optional admission applicants in Study 2 have a significantly lower high school grade point average than standard admission applicants. However, the effect size is extremely small (F = 14.30, p = .00, ηp^2 = .003).

Table 4.5:

Descriptive Statistics on Preadmission Academic Performance of Study 1 Participants

	Test-Optional Applicants			Standard Applicants		
Admission Variable	N	Mean	Standard Deviation	N	Mean	Standard Deviation
High School Grade Point Average	879	3.41	.349	3,974	3.53	.364
SAT Total	446	954.4	116.5	3,896	1,167.0	146.6
SAT Verbal	395	473.5	68.4	3,577	571.8	82.0
SAT Quantitative	395	476.0	69.5	3,577	584.0	82.5
SAT Writing	395	468.2	72.0	3,575	562.7	86.5

Table 4.6:

Descriptive Statistics on Preadmission Academic Performance of Study 2 Participants

	Test-Optional Applicants			Standard Applicants		
Admission Variable	N	Mean	Standard Deviation	N	Mean	Standard Deviation
High School Grade Point Average	717	3.52	.291	4,445	3.56	.341
SAT Total	318	952.9	129.0	4,356	1,183.9	152.3
SAT Verbal	279	467.2	71.9	3,886	578.1	82.9
SAT Quantitative	279	474.5	76.7	3,886	590.7	85.1
SAT Writing	278	465.4	71.9	3,885	568.9	85.7

As is consistent with the literature (e.g., Hiss & Franks, 2014), test-optional admission applicants are more likely to be of lower socioeconomic status (as approximated by the percentage of participants who are Pell Grant recipients) than standard applicants. Specifically, in Study 1, Pell Grant recipients accounted for 44.1% of test-optional participants whereas 25.6% of standard applicants were Pell Grant recipients. Similarly, in Study 2, 42.4% of test-optional applicants were recipients of Pell Grants whereas 24.8% of standard applicants received Pell Grants. These findings are presented in Tables 4.7 and 4.8.

Table 4.7:

Pell Grant Recipients as a Percentage of Participants by Applicant Type

Study 1				Study 2			
Test-Optional Applicants		Standard Applicants		Test-Optional Applicants		Standard Applicants	
N	Percentage	N	Percentage	N	Percentage	N	Percentage
391	44.1	1,034	25.6	307	42.4	1,118	24.8

Additional context is provided in Appendix E which presents descriptive statistics on the predictor variables, which include the aggregate score of the motivational-developmental dimensions (test-optional essay rating score) and traditional admission variables (SAT Total, SAT Verbal, SAT Quantitative, SAT Writing, and high school grade point average). Appendix E shows that the data were normally distributed with acceptable skewness and kurtosis values for the predictor variables as the test statistic for skewness and kurtosis values are greater than -1 and less than 1 (Bulmer, 1979). This suggests that the distribution of predictor variables is approximately symmetric.

Similarly, Appendix F presents descriptive statistics on the outcome variables, cumulative undergraduate grade point average and retention. The descriptive statistics identified that the distribution of outcomes variables is highly negatively-skewed. This suggests there is a greater distribution of values that are less than the mean value for the given variable. However, all outcome variables have acceptable standard error of kurtosis as the test statistic for kurtosis values are greater than -1 and less than 1 (Bulmer, 1979).

Research Question 1: Predictive Validity of Test-Optional Essay Rating Score

Table 4.9 and Table 4.10 present Pearson product-moment correlation coefficients computed to investigate whether the test-optional essay rating score is a significant predictor of student success in college as measured by cumulative undergraduate grade point average and retention. To provide additional context regarding the predictive validity of motivational-developmental dimensions, these tables include the correlation between the outcome variables (cumulative undergraduate grade point average and retention) and several predictor variables frequently considered in college admissions (high school grade point average, SAT Total, SAT Verbal, SAT Quantitative, SAT Writing, and Expected Family Contribution [EFC]). It should be noted that although EFC was included in this analysis, the participating institution practices "need-blind" admissions in which applicants' demonstrated ability to pay for tuition is not considered for the purpose of admission decision-making. Additionally, these correlations coefficients were computed only for test-optional admission applicants.

Table 4.8:

Correlations between Predictor and Outcome Variables in Study 1

Predictor Variable	N	Cumulative Undergraduate Grade Point Average	Semesters Retained
Test-Optional Essay Rating Score	889	.114**	.075*
SAT Total	4,342	.290**	.083**
SAT Verbal	3,972	.257**	.047**
SAT Quantitative	3,972	.233**	.092**
SAT Writing	3,970	.298**	.074**
High School Grade Point Average	4,853	.317**	.074**
Expected Family Contribution	4,264	.101**	.058**

Note. ** $p < .01$ * $p < .05$

Table 4.9:

Correlations between Predictor and Outcome Variables in Study 2

Predictor Variable	N	Cumulative Undergraduate Grade Point Average	Semesters Retained
Test-Optional Essay Rating Score	734	.073*	.101**
SAT Total	4,674	.255**	.051**
SAT Verbal	4,165	.217**	.018
SAT Quantitative	4,165	.188**	.056**
SAT Writing	4,163	.283**	.056**
High School Grade Point Average	5,162	.350**	.037**
Expected Family Contribution	4,442	.074**	.031*

Note. * $p < .05$ ** $p < .01$

The results of these analyses revealed that there was a weak relationship between the test-optional essay rating score variable and cumulative undergraduate grade point average, given Pearson's (1895)[3] conventions to interpret the strength of a correlation as measured by the correlation coefficient *r*. This finding was consistent in both Study 1 ($r = .11$, $p < .01$) and Study 2 ($r = .07$, $p < .05$). Similarly, there was a weak relationship between the test-optional essay rating score and retention in Study 1 ($r = .08$, $p < .05$) and Study 2 ($r = .10$, $p < .01$).

[3] Accepted interpretations indicate that correlation coefficients whose magnitude are between 0.9 and 1.0 indicate variables are very highly correlated. Correlation coefficients between 0.7 and 0.9 indicate that variables are highly correlated. Correlation coefficients between 0.5 and 0.7 indicate that variables are moderately correlated. Correlation coefficients between 0.3 and 0.5 indicate that variables have a low correlation. Correlation coefficients less than 0.3 indicate that variables have little if any (linear) correlation.

According to Sink and Mvududu (2010), larger samples are associated with greater statistical power and reduced sampling error. However, statistically significant effects with trivial effect sizes are more readily identified. Therefore, the aforementioned statistical significance between the predictor and outcome variables can be attributed to the large sample sizes in Study 1 (N = 889) and Study 2 (N = 734). Given these findings, a change in the predictor variable does not strongly correlate with a change in an outcome variable. In other words, the test-optional essay rating score is the weakest predictor of cumulative undergraduate grade point average among the other predictor variables examined in the current study such as high school grade point average and SAT scores.

The results also indicated that there was a weak, positive correlation between the criteria frequently considered in college admissions (high school grade point average, SAT Total, SAT Verbal, SAT Quantitative, SAT Writing, and EFC) and the outcome variables investigated in the current study (cumulative undergraduate grade point average and retention). As is consistent with the findings of previous studies (e.g., Atkinson & Geiser, 2009), high school grade point average had the strongest correlation to cumulative undergraduate grade point average among the predictor variables examined in Study 1 ($r = .32$, $p < .01$) and Study 2 ($r = .35$, $p < .01$). It should be noted that among all of the predictor variables examined in Study 2, the test-optional essay rating score had the strongest correlation ($r = .10$, $p < .01$) to the number of semesters in which participants were retained at the participating institution. However, a weak relationship exists between these variables.

Incremental prediction of test-optional essay rating score. Tables 4.10 through 4.13 include the results of a hierarchical linear regression conducted to determine whether motivational-developmental constructs are incrementally predictive of cumulative undergraduate grade point average and retention given the variance accounted for by traditional admission criteria (high school grade point average, SAT Verbal, SAT Quantitative, and SAT Writing). As shown in these tables, scores on a measure of motivational-developmental dimensions provide marginal value in predicting college outcomes beyond the variance accounted for by the variables most frequently considered in college admissions.

The addition of the test-optional essay rating score variable to the regression model explained 0.8% of the variance in cumulative undergraduate grade point average in Study 1 and 0.6% in Study 2. The addition of the test-optional essay rating score variable to the regression model explained 1.24% of the variance in retention in Study 1 and 1.33% in Study 2.

Table 4.10:

Hierarchical Linear Regression Model for Undergraduate Grade Point Average in Study 1

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.342[a]	.117	.108	.657
2	.353[b]	.125	.114	.655

Note. a. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing
b. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing, Test Optional Essay Rating Score

Table 4.11:

Hierarchical Linear Regression Model for Undergraduate Grade Point Average in Study 2

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.422[a]	.178	.166	.638
2	.429[b]	.184	.170	.637

Note. a. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing
b. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing, Test Optional Essay Rating Score

Table 4.12:

Hierarchical Linear Regression Model for Retention in Study 1

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.075[a]	.006	-.004	.805
2	.112[b]	.013	.000	.803

Note. a. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing
b. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing, Test Optional Essay Rating Score

Table 4.13:

Hierarchical Linear Regression Model for Retention in Study 2

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.086[a]	.007	-.007	.265
2	.119[b]	.014	-.004	.265

Note. a. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing
b. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing, Test Optional Essay Rating Score

Diagnostic accuracy of test-optional essay rating score. Receiver operating characteristic (ROC) analyses were conducted to determine at what test-optional essay rating cut score students are predicted to persist in college. First- and second-year retention was examined for Study 1 and first-year retention was examined for Study 2. As indicated in Table 4.14, 93.8% of test-optional participants in Study 1 were retained during their first-year and 84.1% through their second-year of enrollment. In Study 2, 93.7% of test-optional participants were retained during their first-year of college. These data are evidence of a positive actual state of student retention at the participating institution. In other words, a greater percentage of participants retained in college than was predicted by the median test-optional essay rating score.

Table 4.14:

Summary of Test-Optional Applicant Retention in Studies 1 and 2

Model	N	Retained First-Year	Not Retained First-Year	Retained Second-Year	Not Retained Second-Year
Study 1	889	834	55	748	141
Study 2	734	688	46	N/A	N/A

As presented in Tables 4.15 and 4.16 and Figures 4.1 and 4.2, the area under the curve (AUC) in Study 1 (AUC = .54) and Study 2 (AUC = .62) further supports this finding, given the interpretation of AUC as advanced by Fawcett (2006) and Swets, Dawes, and Monahan (2000)[4]. The AUC was not statistically significant in Study 1 (p = .098), but was significant in Study 2 (p = .007). The 95% confidence intervals do not include the null hypothesis of 0.50. The ROC

[4] Swets, Dawes, and Monahan (2000) have offered benchmarks for gauging AUCs, suggesting that values ≥ 0.9 are ‘‘excellent,’’ ≥ 0.80 ‘‘good,’’ ≥ 0.70 ‘‘fair,’’ and < 0.70 “poor.”

curve coordinates in Appendix G were used to determine that the test-optional essay rating score that maximizes sensitivity and specificity values was 25.565 in Study 1 and 22.44 in Study 2. In Study 1, the test-optional essay rating cutoff score had a sensitivity of .586 and 1 - specificity of .489. In Study 2, the cutoff score had a sensitivity of .593 and 1 – specificity of .413. Given these findings, the test-optional essay rating score did not predict retention in a way that was meaningfully different from chance. Therefore, this variable does not serve as an accurate binary classifier or diagnostic predictor of retention and attrition for test-optional participants.

Table 4.15:

Area Under the Curve (AUC) for Study 1

			95% Confidence Interval	
Area	Std. Error[a]	Asymptotic Sig.[b]	Lower Bound	Upper Bound
.544	.027	.098	.490	.598

Note. a. Under the nonparametric assumption
b. Null hypothesis: true area = 0.5

Table 4.16:

Area Under the Curve (AUC) for Study 2

			95% Confidence Interval	
Area	Std. Error[a]	Asymptotic Sig.[b]	Lower Bound	Upper Bound
.618	.040	.007	.539	.696

Note. a. Under the nonparametric assumption
b. Null hypothesis: true area = 0.5

Figure 4.1:

ROC Curve for Test-Optional Applicant Retention in Study 1

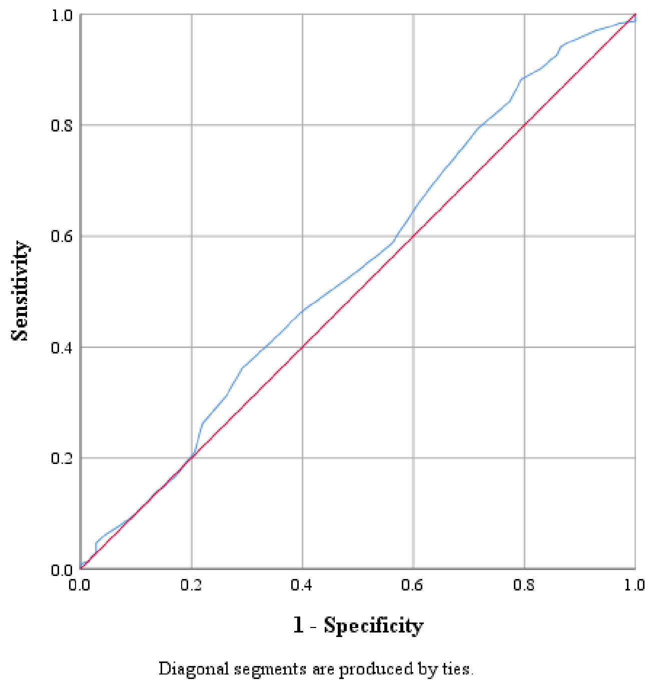
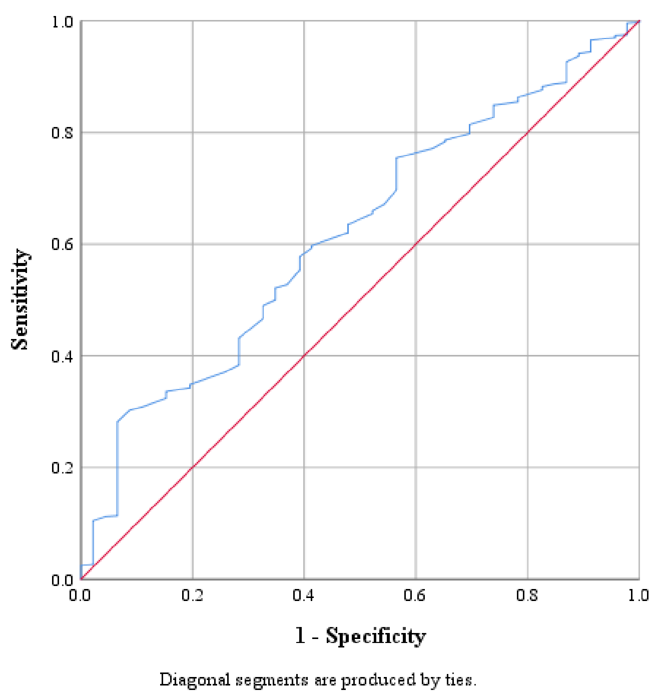


Figure 4.2:

ROC Curve for Test-Optional Applicant Retention in Study 2



Contribution of Specific Motivational-Developmental Dimensions to Prediction of Student Success. Table 4.17 presents the findings of a Pearson correlation analysis conducted to determine the contribution, if any, of the underlying motivational-developmental dimensions which comprise the test-optional essay rating score variable to the prediction of college student success as measured by cumulative undergraduate grade point average and retention. However, given the weak relationship between the test-optional essay rating score and the outcomes variables, the value of this secondary analysis was minimal.

Among the motivational-developmental dimensions examined in the current study, self-awareness ($r = .12$, $p < .01$), coping ($r = .11$, $p < .01$), and self-set goals ($r = .11$, $p < .01$) were the greatest predictors of cumulative undergraduate grade point average in Study 1. Essays 3 and 4 assessed a combination of these dimensions in Study 1. Attributions ($r = .11$, $p < .01$) was the greatest predictor of cumulative undergraduate grade point average in Study 2.

Coping ($r = .08$, $p < .05$), self-awareness ($r = .07$, $p < .05$), and self-authorship ($r = .07$, $p < .05$) were the greatest predictors of retention in Study 1. Essays 3 and 4 assessed a combination of these dimensions in Study 1. Self-authorship ($r = .11$, $p < .01$) and self-awareness ($r = .10$, $p < .01$) were the greatest predictors of retention in Study 2. Essays 3 and 4 assessed a combination of these dimensions in Study 2.

Table 4.17:

Correlation between Specific Motivational-Developmental Dimensions and Outcome Variables

	Study 1		Study 2	
Dimension	Cumulative Undergraduate Grade Point Average	Semesters Retained	Cumulative Undergraduate Grade Point Average	Semesters Retained
Attributions	.090**	.061	.114**	.070*
Self-Authorship	.087**	.068*	.057	.110**
Self-Awareness	.120**	.071*	.062	.100**
Coping	.111**	.076*	.005	.013
Future Time Perspective	N/A	N/A	.044	.044
Relevant Experiences	.037	.029	N/A	N/A
Self-Concept	.054	.026	.002	.035
Self-Set Goals	.109**	.046	.019	.035

Note. * $p < .05$ ** $p < .01$

Research Question 2: Differential Prediction of Test-Optional Essay Rating Score

As shown in Tables 4.18 and 4.19, a moderator analysis revealed that the correlations between the test-optional essay rating score and cumulative undergraduate grade point average and semesters retained vary by gender and race. For Study 1, the test-optional essay rating score correlated significantly ($r = .13$, $p < .01$) with cumulative undergraduate grade point average for females, but not for male participants ($r = .08$). Conversely, the correlation for retention showed the opposite pattern in which the test-optional essay rating score correlated significantly for males ($r = .14$, $p < .05$), but not for female participants ($r = .04$). However, in each case, the difference between the correlations was not statistically significant. For Study 2, the test-optional essay rating score correlated significantly for retention for females (r = .13, $p < .01$), but not for males (r = .03). As before, the correlations were not significantly different. None of the correlations by race were significant.

Table 4.18:

Correlations between Test-Optional Essay Rating Score and Outcome Variables by Gender

	Study 1		Study 2	
Reported Gender	Cumulative Undergraduate Grade Point Average	Semesters Retained	Cumulative Undergraduate Grade Point Average	Semesters Retained
Females	.128**	.040	.058	.128**
Males	.081	.136*	.091	.034

Note. * $p < .05$ ** $p < .01$

Table 4.19:

Correlations between Test-Optional Essay Rating Score and Outcome Variables by Race

Race	Study 1 Cumulative Undergraduate Grade Point Average	Study 1 Semesters Retained	Study 2 Cumulative Undergraduate Grade Point Average	Study 2 Semesters Retained
African American	.064	.124	.144	.117
Asian	.108	.112	-.041	.087
Hispanic	-.098	-.067	-.012	.111
White	.066	.026	.022	.085
Other	.188	.172	.165	.058

Secondary Research Questions

The results of analyses that respond to the secondary research questions lend further insight into the selected measure of motivational-developmental dimensions and the potential of the NSQ as a predictor of college student success.

Meeting or exceeding predicted cumulative undergraduate grade point average. A secondary research question investigated whether test-optional admission applicants who enroll in college are exceeding expectations in terms of academic performance as measured by cumulative undergraduate graduate point average. Specifically, since the SAT is known to predict first-year grade point average, and since the test-optional applicants have, by definition, lower SAT scores, their cumulative undergraduate grade point average would be expected to be lower than standard admission applicants who submitted SAT scores. To demonstrate if this is correct, the cumulative undergraduate grade point averages of the test-optional and standard applicants were compared. Tables 4.20 and 4.21 present these data.

Table 4.20:

Means and Standard Deviations for Undergraduate Grade Point Average for Study 1

Semester	Application Type	N	Mean	Standard Deviation
Semester 1	Test-Optional	746	2.91	.677
	Standard	3,520	3.23	.644
	Total	4,266	3.17	.661
Semester 2	Test-Optional	746	2.95	.597
	Standard	3,520	3.22	.591
	Total	4,266	3.17	.601
Semester 3	Test-Optional	746	2.99	.540
	Standard	3,520	3.23	.548
	Total	4,266	3.19	.554
Semester 4	Test-Optional	746	3.01	.530
	Standard	3,520	3.24	.534
	Total	4,266	3.19	.540

Table 4.21:

Means and Standard Deviations for Undergraduate Grade Point Average for Study 2

Semester	Application Type	N	Mean	Standard Deviation
Semester 1	Test-Optional	749	2.99	.542
	Standard	3,527	3.23	.548
	Total	4,276	3.19	.554
Semester 2	Test-Optional	749	3.01	.532
	Standard	3,527	3.24	.534
	Total	4,276	3.20	.541

Repeated measures analysis of variance (ANOVAs) were computed on the data for Study 1 and Study 2. In both cases, there was a significant main effect for group and a significant interaction. Plots of the data for the two studies are presented in Figures 4.3 and 4.4. As shown in these figures, test-optional participants have lower cumulative undergraduate grade point averages in both Study 1 and Study 2. To determine if the test-optional participants are performing at a higher level than would be predicted, the regression model presented above was used to predict cumulative undergraduate grade point average. These predicted scores were then compared to participants' actual cumulative undergraduate grade point average. The mean predicted cumulative undergraduate grade point average across both studies for the test-optional participants was 2.92; the actual cumulative undergraduate grade point average was 2.93. This difference was not statistically significant.

Figure 4.3:

Undergraduate Grade Point Average for Study 1

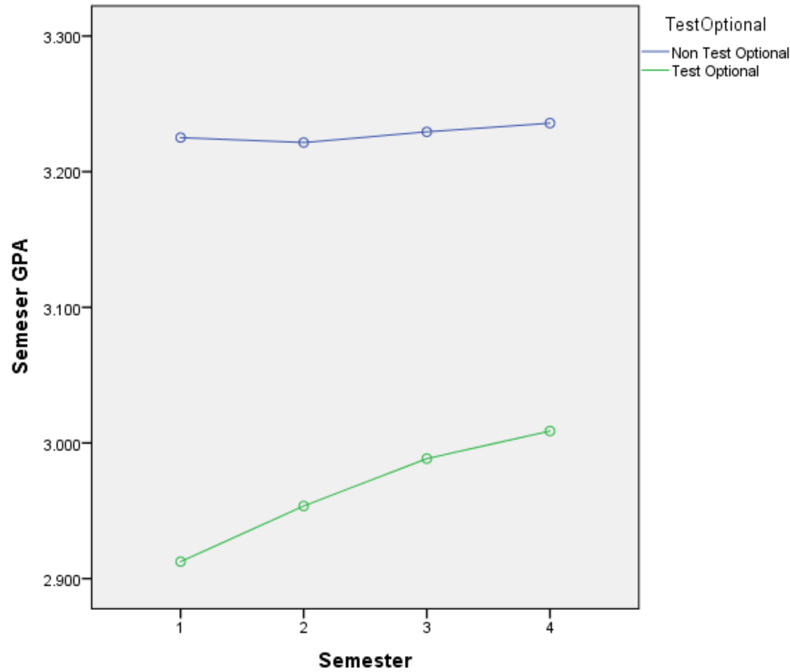
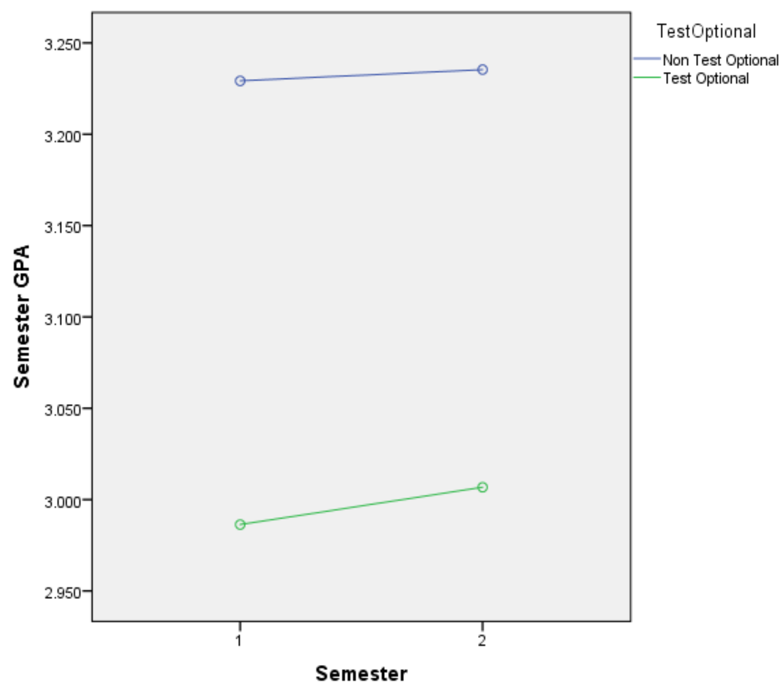


Figure 4.4:

Undergraduate Grade Point Average for Study 2



Predictive validity of New Student Questionnaire (NSQ) items. Pearson correlations were computed relating the NSQ motivational items (see Appendix B) with cumulative undergraduate grade point average and retention. The correlations for Study 1 are presented in Table 4.22 and Table 4.23 for Study 2. The results of these analyses revealed that there is a weak relationship between the NSQ motivational items and the outcome variables, cumulative undergraduate grade point average and retention. This finding is consistent for both Study 1 and Study 2. The statistical significance between the NSQ items and the outcome variables can be attributed to the large number of participants in Study 1 (N = 4,733) and Study 2 (N = 4,985). These analyses included the available NSQ data for both test-optional and standard admission applicants.

It is useful to note that there was a small, but statistically significant positive correlation ($r = .17$, $p < .01$) between cumulative undergraduate grade point average and participants' perception that their teachers considered them one of the harder workers in their high school classes. Conversely, there was a small, but statistically significant negative correlation ($r = -.03$, $p < .01$) between perceived self-confidence and cumulative undergraduate grade point average. This suggests that the greater a participant's self-confidence, the lower their cumulative undergraduate grade point average.

Table 4.22:

Correlations between NSQ Items and Outcome Variables for Study 1

NSQ Item	N	Cumulative Undergraduate Grade Point Average	Semesters Retained
Hard Work	4,733	.154**	.039**
Difficult Plan	4,716	.114**	.031*
Enjoy Study	4,730	.003	-.015
Time Management	4,739	.067**	.024
Confident	4,711	-.042**	.014
Give Up	4,708	.034*	.019
Organized	4,725	.128**	.049**
Independent	4,728	-.020	-.045**

Note. * $p < .05$ ** $p < .01$

Table 4.23:

Correlations between NSQ Items and Outcome Variables for Study 2

NSQ Item	N	Cumulative Undergraduate Grade Point Average	Semesters Retained
Hard Work	4,985	.186**	.022
Difficult Plan	4,967	.107**	.026
Enjoy Study	4,969	.027	-.011
Time Management	4,980	.095**	.034*
Confident	4,969	-.017	.006
Give Up	4,958	.058**	.014
Organized	4,969	.145**	.037**
Independent	4,973	-.007	-.035*

Note. * $p < .05$ ** $p < .01$

Tables 4.24 through 4.27 include the results of ordinary least squares regressions conducted on the data presented in Tables 4.22 and 4.23. These analyses were conducted to determine whether the NSQ items are incrementally predictive of cumulative undergraduate grade point average and retention. As shown in these tables, the NSQ motivational items in Study 1 account for 4.8% of the variance in cumulative undergraduate grade point average and 0.5% of the variance in retention. The NSQ motivational items in Study 2 account for 5.9% of the variance in cumulative undergraduate grade point average and 0.2% of the variance in retention.

Table 4.24:

NSQ Items Linear Regression Model for Undergraduate Grade Point Average in Study 1

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.220[a]	.048	.046	.710

Note. a. Predictors: (Constant), Independent, Difficult Plan, Enjoy Study, Confident, Hard Work, Give Up, Time Management, Organized

Table 4.25:

NSQ Items Linear Regression Model for Undergraduate Grade Point Average in Study 2

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.246[a]	.061	.059	.709

Note. a. Predictors: (Constant), Independent, Give Up, Enjoy Study, Hard Work, Confident, Difficult Plan, Time Management, Organized

Table 4.26:

NSQ Items Linear Regression Model for Retention in Study 1

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.083[a]	.007	.005	.756

Note. a. Predictors: (Constant), Independent, Give Up, Enjoy Study, Hard Work, Confident, Time Management, Difficult Plan, Organized

Table 4.27:

NSQ Items Linear Regression Model for Retention in Study 2

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.064[a]	.004	.002	.206

Note. a. Predictors: (Constant), Independent, Give Up, Hard Work, Enjoy Study, Confident, Difficult Plan, Time Management, Organized

The aforementioned NSQ items were added to a multiple regression model to examine their incremental value in predicting cumulative undergraduate grade point average and retention given the percentage of the variance accounted for by traditional admission criteria (high school grade point average, SAT Verbal, SAT Quantitative, SAT Writing) and the test-optional essay rating score variable. Tables 4.28 through 4.31 present the results of these regression models. As shown in these tables, the NSQ motivational items provide nominal value to the prediction of college outcomes beyond what is accounted for by the aforementioned predictor variables.

The addition of the NSQ motivational items to the regression model explained 5.5% of the variance in cumulative undergraduate grade point average in Study 1 and 2.0% in Study 2. The addition of the NSQ motivational items to the regression model explained 2.2% of the variance in retention in Study 1 and 3.7% in Study 2.

Table 4.28:

NSQ Items Hierarchical Linear Regression Model for Undergraduate Grade Point Average in Study 1

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.342[a]	.117	.108	.657
2	.353[b]	.125	.114	.655
3	.424[c]	.180	.150	.629

Note. a. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing
b. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing, Test Optional Essay Rating Score
c. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing, Test Optional Essay Rating Score, Independent, Give Up, Hard Work, Enjoy Study, Confident, Difficult Plan, Time Management, Organized

Table 4.29:

NSQ Items Hierarchical Linear Regression Model for Undergraduate Grade Point Average in Study 2

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.422[a]	.178	.166	.638
2	.429[b]	.184	.170	.637
3	.451[c]	.204	.162	.654

Note. a. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing
b. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing, Test Optional Essay Rating Score
c. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing, Test Optional Essay Rating Score, Independent, Give Up, Hard Work, Enjoy Study, Confident, Difficult Plan, Time Management, Organized

Table 4.30:

NSQ Items Hierarchical Linear Regression Model for Retention in Study 1

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.075[a]	.006	-.004	.805
2	.112[b]	.013	.000	.803
3	.187[c]	.035	.000	.779

Note. a. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing
b. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing, Test Optional Essay Rating Score
c. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing, Test Optional Essay Rating Score, Independent, Give Up, Hard Work, Enjoy Study, Confident, Difficult Plan, Time Management, Organized

Table 4.31:

NSQ Items Hierarchical Linear Regression Model for Retention in Study 2

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.086[a]	.007	-.007	.265
2	.119[b]	.014	-.004	.265
3	.225[c]	.051	.001	.262

Note. a. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing
b. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing, Test Optional Essay Rating Score
c. Predictors: (Constant), High school grade point average, SAT Verbal, SAT Quantitative, SAT Writing, Test Optional Essay Rating Score, Independent, Give Up, Hard Work, Enjoy Study, Confident, Difficult Plan, Time Management, Organized

Interrater Reliability

Light's kappa (Light, 1971) was computed to determine the agreement between readers on each of the motivational-developmental dimensions assessed in the test-optional essays.

Light's kappa is used to measure interrater reliability when there is not a fixed number of raters. Light's kappa is computed by calculating Cohen's kappa and averaging the values manually across all rater pairs. Table 4.32 presents the results of the Light's kappa measure of agreement across raters and the motivational-developmental dimensions assessed in each test-optional admission essay. Given the interpretations of kappa values as suggested by Cohen (1960)[5], there was primarily slight agreement between raters across the motivational-developmental dimensions. This finding is consistent for both Study 1 and 2.

Table 4.32:

Interrater Reliability Analysis

	Study 1	Study 2
Motivational-Developmental Dimension	Light's Kappa Measure of Agreement	Light's Kappa Measure of Agreement
Attributions	.191	.128
Authorship	.218	.187
Awareness	.205	.125
Coping	.132	.126
Future Time Perspective	N/A	.183
Relevant Experiences	.238	N/A
Self-Concept	.155	.208
Self-Set Goals	.146	.112

[5] Cohen suggested the kappa result be interpreted as follows: ≤ 0 as indicating no agreement and 0.01-0.20 as none to slight, 0.21-0.40 as fair, 0.41-0.60 as moderate, 0.61-0.80 as substantial, and 0.81-1.00 as almost perfect agreement (Cohen, 1960; McHugh, 2012).

CHAPTER 5

DISCUSSION

Research has established the need to identify and validate additional predictors of college student success in order to account for a greater percentage of the variance in college outcomes and attain the broader goals of postsecondary education. Despite these understandings, there is much to consider regarding the predictive potential of motivational-developmental dimensions and their use as criteria for the evaluation and selection of admission applicants. Motivational-developmental constructs are under-explored in terms of their predictive potential in measuring admission applicants' ability to persevere despite the challenges presented by postsecondary participation.

The current study investigated the predictive validity of a measure of motivational-developmental dimensions in predicting college student success as defined by cumulative undergraduate grade point average and retention. The current study aimed further to contribute understandings to the body of research on test-optional admissions and the various enrollment management models employed by postsecondary institutions. Furthermore, the findings of the current study are a contribution to the scholarly and applied efforts to account for a greater percentage of the variance in college student success, as measured in terms of academic performance and persistence.

This chapter summarizes the main findings of the current study. It also addresses limitations of the research conducted, provides implications for the use of motivational-developmental dimensions in the context of college admissions, and offers directions for future research. The ensuing discussion is an interpretation of the findings in an attempt to answer the research questions presented in Chapter 1.

Summary of Findings

A primary objective of the current study was to determine the validity of a measure of motivational-developmental dimensions as a predictor of the academic achievement and persistence of college students. The current study revealed that there is a not a strong correlation between the test-optional essay rating score and cumulative undergraduate grade point average or retention, particularly in comparison to the relationship between these outcome variables and the predictors most commonly considered in college admissions (high school grade point average, SAT Verbal, SAT Quantitative, and SAT Writing). Despite this finding, the test-optional essay rating score contributed nominal value in predicting academic achievement and persistence of the participants beyond the explanation provided by traditional admissions criteria. However, its contribution was not of statistical or practical significance, particularly given the proportion of the variance that remains unexplained by the combination of the test-optional essay rating score and the predictor variables traditionally used in college admissions.

Given the well-documented implications of leaving college without obtaining a degree or credential (e.g., Pascarella & Terenzini, 1991), there is applied interest in validating predictors of college student persistence prior to their participation in postsecondary education. Thus, a primary purpose of the current study was to evaluate the predictive accuracy of the test-optional essay rating score as a classifier of student retention. The current study found that the test-optional essay rating score is not an accurate diagnostic predictor of retention and attrition for those who apply for admission using a test-optional admission policy and subsequently enroll in college.

The current study also intended to determine if specific motivational-developmental dimensions contribute most significantly to the prediction of cumulative undergraduate grade point average and retention. This aim is of particular importance given that the current study examined a subset of the available motivational-developmental dimensions. Therefore, the identification of specific motivational-developmental dimensions with the greatest correlation to college student success informs future efforts to validate similar predictors. Among the variables examined in the current study, awareness, coping, and self-set goals for cumulative undergraduate grade point average and casual attributions for retention have the greatest promise for predicting college student success. However, there was a weak correlation between these dimensions and the outcome variables, which signals the need to investigate the validity of other motivational and developmental constructs as predictors of college student success. This finding is consistent with the weak relationship between the test-optional essay rating score and outcome variables. This finding can be attributed to the fact that the test-optional essay rating score is inclusive of the individual motivational-developmental dimension values. Nonetheless, this finding is an important contribution to the understanding of how specific motivational-developmental constructs may relate to indicators of student success.

Another primary objective of the current study was to determine whether the test-optional essay rating score is differentially predictive of the cumulative undergraduate grade point average and retention for participants of racial and gender subgroups. In Study 1, the test-optional essay rating score had a weak, but statistically significant correlation to cumulative undergraduate grade point average for female participants ($r = .13$, $p < .01$). However, there was not a statistically significant relationship between these variables in Study 2. The test-optional essay rating score had a weak, but statistically significant correlation to retention for male participants in Study 1 ($r = .136$, $p < .05$) and for female participants in Study 2 ($r = .128$, $p < .01$). In all cases, the difference between the correlations was not statistically significant.

Additionally, none of the correlations by racial subgroup were significant. Given these findings, there is inconclusive evidence regarding whether the test-optional essay rating score differentially predicts cumulative undergraduate grade point average. However, there is evidence to suggest that the test-optional essay rating score does not differentially predict cumulative undergraduate grade point average and retention across racial subgroups.

A secondary objective of the current study was to investigate whether test-optional applicants who enroll in college attain a greater or lesser cumulative undergraduate point average than is predicted by their SAT scores. The current study found that there is not a statistically significant difference in the predicted and observed cumulative undergraduate grade point average attained by test-optional applicants. Given this finding, there is evidence that test-optional applicants do not outperform the expected cumulative undergraduate point average as predicted by their SAT scores.

Additionally, the current study examined whether the New Student Questionnaire (NSQ) items that measure student motivation are significant predictors of cumulative undergraduate grade point average and retention. On average, the NSQ items have stronger correlations to cumulative undergraduate grade point average compared to the test-optional essay rating score. Conversely, the test-optional essay rating score is a stronger predictor of retention than the NSQ items. However, there are not statistically significant differences between the correlations.

In combination, the predictor variables examined in the current study (test-optional essay rating score, high school grade point average, SAT Verbal, SAT Quantitative, SAT Writing, and NSQ motivational items) accounted for approximately 19.2% of the variance in cumulative undergraduate grade point average and approximately 4.3% of the variance in retention. Specifically, the test-optional essay rating score accounted for approximately 0.9% of the variance in cumulative undergraduate grade point average and approximately 0.8% of the variance in retention. As discussed in Chapter 2, previous studies have explained a similar percentage of the variance in college student success. For example, Geiser and Santelices (2007) successfully explained 24.5% of the variance in first-year grade point average using a regression model that included high school grade point average, SAT scores, parental education level, family income, and school Academic Performance Index (API) rank. However, similar to the current study, their findings lack external validity as the study was conducted exclusively at the University of California. Despite the understandings gleaned from the current study and those conducted previously, a substantial percentage of the variance in college success remains unexplained by the variables examined in the current study.

Discussion of Findings

Since the first administration of the SAT in 1901, there has been applied interest in identifying variables that reliably and accurately predict college outcomes. The current study sought to serve as a contribution to these efforts, particularly given the increased national emphasis on expanding postsecondary participation and completion. This study also aimed to generate new understandings to benefit the participating institution and other institutions that seek new methodologies for evaluating and selecting college admission applicants.

The current study revealed a weak relationship between the test-optional essay rating score and the outcome variables (cumulative undergraduate grade point average and retention). This finding raises questions regarding the reliability, construct validity, and predictive validity of the test-optional essay rating score as an aggregate variable that is reflective of the motivational-developmental dimensions examined in the current study. It is also important to consider whether motivational-developmental dimensions more accurately predict the broad and variable definition of college student success, as compared to the relatively narrow definition employed by the current study and in the majority of the literature.

Reliability of the measure. Given the prominence of college admissions as a gateway to postsecondary participation, there is applied interest in examining whether the prediction of college student success varies across groups of college applicants who differ on demographic variables such as race, gender, and socioeconomic status (Young & Kobrin, 2001). The reliability (in terms of both the integrity of data and the consistency of measurement) of admission criteria is prerequisite for their validity as predictors of the future performance of college admission applicants. Therefore, these predictors have the greatest potential when there is evidence of equivalent means and variances across subgroups on the criterion variable(s). As noted by the *Standards for Educational and Psychological Testing*,

> Correlation coefficients provide inadequate evidence for or against a differential prediction hypothesis if groups or treatments are found to have unequal means and variances on the test and the criterion. It is particularly important in the context of testing for high-stakes purposes that test developers and/or users examine differential prediction and avoid the use of correlation coefficients in situations where groups or treatments result in unequal means or variances on the test and criterion. (AERA, APA, & NCME, 2014, p. 66)

Findings of the current study suggest that the test-optional essay rating score may reduce the reliance on standardized test scores for admission decision-making by accounting for some of the variance in student success, particularly for female participants. However, given the small percentage of the variance accounted for by the test-optional essay rating score, further investigation is required to determine if this variable is differentially predictive of student success across ethnic and gender subgroups.

Construct validity of the measure. Although the current study focused on the predictive validity of motivational-developmental dimensions, it is important to consider issues of construct validity that affect the predictive potential derived from the test-optional essay rating score. A measure has construct validity if it measures what it intends to measure. Thus, it is important to consider whether the test-optional essay rating score serves its intended purpose as an evaluative measure of the motivation and development of college admission applicants in order to make a prediction about their ability to achieve academically and persevere despite the intellectual and social challenges presented by postsecondary education. For example, the Future Time Perspective dimension requires applicants to consider how their previous experiences may support their success in the future. This dimension may reflect applicants' human capital (i.e., knowledge of what to expect as a college student) as opposed to their ability to articulate the connections between their previous and future experiences as they relate to postsecondary participation. Similarly, the essays that measure the Relevant Experiences dimension required test-optional applicants to make connections between their previous experiences and those that they may experience as a college student.

The current study found minimal convergent evidence to suggest that the test-optional essay rating score is strongly associated with the indicators of college student success that it intends to measure. This weak association may be both a function of limited content-related validity and the fact that the outcome variables reflect a narrow view of college student success.

Definition of college student success. Although the majority of predictive validity studies use first-year grades as a proxy for success in college, long-term success in college is an equally important criterion to many institutions (Burton & Ramist, 2001; Harackiewicz, Barron, Tauer, & Elliot, 2002). As hypothesized by Spence and Helmreich (1983), achievement goals predict outcomes in the short term whereas general measures of achievement motivation may prove to be stronger predictors of academic performance when aggregated over a longer period.

This hypothesis is consistent with the findings of the current study, as motivational-developmental dimensions are weak predictors of short-term academic performance. Therefore, further research should determine whether motivational-developmental dimensions serve as valid predictors of long-term indicators of college success such as post-college employment and earnings, graduate school acceptance, and civic participation.

The usefulness of motivational-developmental dimensions as admissions criteria depends on institutional objectives and values. For example, the intended use of the test-optional essay rating score as a diagnostic predictor of student retention may depend on admissions selectivity. Presumably, institutions with a low rate of acceptance (i.e., high admissions selectivity) prefer to use a diagnostic measure with a greater rate of false negatives. In other words, institutions that intend to maintain a low rate of acceptance may prefer to deny applicants who would have been retained at the institution had they been admitted. Conversely, institutions with access-based missions and/or those that provide effective academic and student support services may prefer a greater rate of admitted applicants who will have a lesser chance of succeeding and retaining at the institution. Thus, the interpretations of the findings presented in the current study depend on their intended use and alignment with institutional priorities.

Format of the measure. Given the findings presented in Tables 4.22 and 4.23, the dimensions with the strongest correlations to the outcomes variables may more accurately reflect the groupings of motivational-developmental dimensions within each essay question as opposed to inherent relationships between these variables. However, these findings suggest that it may be feasible to reduce the total number of essays by placing greater emphasis on the dimensions with the strongest correlation to the outcome variables. A reduction in the number of essays that test-optional applicants are required to submit may result in an increased rate by which test-optional applicants submit a complete application for admission. As a point of reference, the rate of application completion was 91.5% for test-optional applicants and 97.5% for standard applicants.

As mentioned earlier in this chapter, the current study revealed smaller correlations between the test-optional essay rating score and the outcome variables as compared to the correlations between the NSQ motivational items and the same outcome variables. Therefore, there is evidence that suggests motivational-developmental dimensions may be more accurately measured using a multiple item inventory as opposed to open-ended essay questions scored by a human reader. The calculation of Light's kappa revealed that there was minimal agreement between readers on each of the motivational-developmental dimensions. A greater level of agreement between items is one of the potential benefits of an inventory as opposed to open-ended essay questions. Additionally, a machine-scored assessment would reduce the time by which the results are computed, allowing institutions to provide applicants with admission decisions more quickly.

Readers of the test-optional essays were trained extensively in order to advance a shared definition of the motivational-developmental dimensions assessed in each test-optional essay question. Despite these trainings, interpretations of kappa values suggest there was slight agreement between the raters. According to Hallgreen (2012),

> In studies using trained coders, it may often be necessary to conduct a considerable amount of training with practice subjects before subjects from the real study are coded. In these cases, it is common to specify an *a priori* level of inter-rater reliability (IRR) that must be achieved before subjects from the real study are rated and to report this in the final study write-up. Commonly, the qualitative ratings for different IRR statistics can be used to assign these cutoff points; for example, a researcher may require all IRR estimates to be at least in the "good" range before coders can rate the real subjects in a study. (p. 4)

There are several recommendations for enhancing this level of agreement including additional training of the raters, calibration of the rubric developed by Kaplan et al. (2017), and the implementation of Automated Essay Scoring Technology (AES). Despite these efforts, there is inherent subjectivity in the evaluation of human subjects based on theoretical constructs in applied contexts. In other words, individual interpretations of motivation and human development affect the consistency of the measurement of motivational-developmental dimensions.

Given the methodology of the current study and the admissions practices of the participating institution, human subjectivity is inherent in the scoring of the test-optional essays. AES technology can mitigate the confounding effects of issues related to interrater and intrarater reliability and thereby enhance the efficacy of the measure. Several studies have assessed the accuracy and reliability of AES systems with respect to writing assessment and have reported high agreement rates between AES systems and human raters (Dikli, 2006).

Limitations

Interpretation of the findings presented in the current study should consider several limitations of this research. As mentioned in Chapter 3, the current study included several delimitations including a restriction in range of the degree level and type of participants, classification and geography of the participating institution, and scope of the selected motivational-developmental constructs.

Participants. A characteristic of the current study that should frame the interpretation of the findings is the relative homogeneity of the participants as compared to the gender and racial demographics of postsecondary participants nationally. Appendix C summarizes the gender and racial representativeness of the study participants compared to national postsecondary participants. The homogeneity of the study participants is particularly apparent given that several races were grouped into a category ("Other") so that quantitative analyses could be feasibly conducted on these subgroups. Therefore, a larger and more diverse sample would afford greater insight into the differential prediction of the test-optional essay rating score and its potential for explaining the variance in college outcomes across student subgroups.

Participants of the current study included undergraduate freshmen enrollees, as is consistent with the majority of the literature. Transfer students, graduate students, and denied admission applicants were excluded from participation. Therefore, findings do not generalize to these student populations, despite the need for clear and consistent evidence across student levels and types. In fact, Sedlacek (2004) posited that non-cognitive variables might demonstrate their greatest potential in the context of professional school (e.g., law school, medical school) admissions in which standardized tests are frequently utilized as predominant criteria.

Long-term success in college. Despite broader definitions of college student success that appear in the literature, the current study measured academic performance and persistence over a relatively short period compared to the average length of a college career. Additionally, the outcome variables examined in the current study do not capture these broader definitions nor the personal characteristics of postsecondary participants and completers. Frequently, the results of postsecondary education are not fully realized until post-graduation (Hossler & Bontrager, 2014). Thus, the understandings offered by the current study reflect a relatively narrow measure of college student success and achievement.

External validity. The current study was conducted at a single institution located in a specific geographic region of the United States. The participating institution reflects a particular institutional classification (i.e., public, urban, comprehensive research university) and therefore is not representative of the range of institutional types that collectively characterize the modern landscape of American higher education. Given this delimitation, the current study lacks external validity, which limits the generalizability of the findings.

Directions for Future Research

The findings of the current study have several implications for the future examination of motivational-developmental dimensions as predictors of college student success. Given the proportion of the variance in college outcomes that remains unexplained, future research should employ a more dynamic approach in assessing the ability of admissions applicants to achieve success and persist in college. In doing so, a multisite, longitudinal investigation is required to capture achievement and persistence over the average length of a college career. Research should also consider the prediction and measurement of outcomes that extend beyond college including, but not limited to, post-college employment and earnings, graduate school acceptance, and civic participation (Burton & Ramist, 2001; Willingham, 1985).

Additionally, future investigation should aim to generate new understandings about the relationship between motivational-developmental dimensions and the behaviors frequently exhibited by successful college students. Such behaviors may include the utilization of academic support resources and services as well as the demonstration of higher levels of social integration, which Tinto (1997) suggests is positively correlated with student retention.

The current study examined a subset of the available motivational-developmental dimensions that may relate to college student success. Although the current study found weak relationships between specific motivational-developmental dimensions and the outcome variables, there remains a need to validate additional non-cognitive predictors of college student success. Thus, future research should focus on investigating the validity of other constructs that are drawn from similar perspectives and domains as those employed in the current study.

Conclusion

The effectiveness of any program or policy is determined by ascertaining whether it is meeting its intended goals. It is useful to consider the goals of the participating institution in implementing a test-optional admission policy. While the goals were never explicitly stated, it is reasonable to assume that the participating institution had similar goals to those as most other institutions that have implemented a test-optional admission policy. Several studies have documented the extent to which test-optional admission policies are effective in attaining these objectives (e.g., Hiss & Franks, 2014; Mulugetta, 2013, Shanley, 2007; Soares, 2012). These goals typically focus on at least four outcomes. Given the assumed bias in SAT scores, the first goal is to implement an admission model that is not subject to this bias. Second, by doing this, the institution aims to diversify its admission cohort. Third, the institution intends to attract a greater volume of admission applicants who subsequently enroll at the institution. Finally, assumedly applicants admitted using this test-optional admission process would be at least as academically successful as the applicants admitted using the standard admission process in which standardized test scores are considered.

The analyses presented in Chapter 4 suggest that the institution has at least partially met the first three goals, but has not met the fourth. Specifically, applicants admitted through the test-optional admission policy are more diverse, there appear to be no significant differences in the test-optional score as a function of ethnic group, and the admission pool has increased. It is important to note that the percentage of test-optional students who are female was far higher than anticipated. In terms of academic success, however, the test-optional students are not performing as well as compared to the students admitted through the standard admission process. Specifically, test-optional applicants have lower cumulative undergraduate grade point averages and a somewhat lower rate of retention as compared to students admitted through the standard admission process.

Given the findings presented in the current study, it is useful to consider the applicants who were not admitted to the participating institution and whether they would have achieved academic success had they been admitted through the test-optional admission policy. This uncertainty is a universal limitation of test validity studies. Nonetheless, the implementation of a different motivational-developmental measure may yield a different cohort of students who might achieve a greater rate of academic success.

Findings of the current study support the use of motivational-developmental dimensions to account for a small percentage of the variance in college outcomes among test-optional admission applicants. However, these dimensions did not have statistically significant correlations with academic achievement indicators as measured by cumulative undergraduate grade point average and retention. Despite these understandings, the need to account for a greater percentage of the variance in college outcomes requires future research to investigate the predictive validity of motivational-developmental dimensions in a wider range of institutional contexts. Additionally, there remains a need to validate predictors that are derived from different non-cognitive domains and are minimally correlated with traditional predictors of college student success.

REFERENCES

Aguinis, H., Culpepper, S. A., & Pierce, C. A. (2016). Differential prediction generalization in college admissions testing. *Journal of Educational Psychology*, *108*(7), 1045.

Allen, J., Robbins, S. B., & Sawyer, R. (2010). Can measuring psychosocial factors promote college success? *Applied Measurement in Education*, *23*(1), 1-22. doi:10.1080/08957340903423503.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). Standards for educational and psychological testing. Washington, DC: National Council on Measurement in Education.

American Psychological Association, Task Force on Socioeconomic Status. (2007). Report of the APA Task Force on Socioeconomic Status. Washington, DC: American Psychological Association.

Ancis, J. R., & Sedlacek, W. E. (1997). Predicting the academic achievement of female students using the SAT and noncognitive variables. *College and University*, *72*, 1–8.

Astin, A. W. (1993). What matters in college?: Four critical years revisited. San Francisco: Jossey-Bass.

Atkinson, R., & Geiser, S. (2009). Reflections on a century of college admissions tests. *Educational Researcher*, *38*, 665-676.

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *Journal of Social and Clinical Psychology*, *4*(3), 359-373.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.). *Self-efficacy beliefs of adolescents*, (Vol. 5., pp. 307-337). Greenwich, CT: Information Age Publishing.

Baum, S., Ma, J., & Payea, K. (2013). Education pays: The benefits of higher education for individuals and society. College Board Trends in Higher Education Series. http://trends.collegeboard.org/sites/default/files/education-pays-2013-full-report.pdf.

Baxter-Magolda, M., Creamer, E., & Meszaros, P. (2010). Development and assessment of self-authorship: Exploring the concept across cultures. Sterling: Stylus Publishing.

Belasco, A. S., Rosinger, K. O., & Hearn, J. C. (2015). The test-optional movement at America's selective liberal arts colleges: A boon for equity or something else? *Educational Evaluation and Policy Analysis*, *37*(2), 206-223.

Bowen, W. G., & Bok, D. (1998). The shape of the river. Princeton, NJ: Princeton University Press.

Bridgeman, B., McCamley-Jenkins, L. & Ervin, N. (2000). Predictions of freshman grade point average from the revised and recentered SAT-1: Reasoning test. (RR 2000-1). New York, NY: College Board.

Brooks, S. (2014). Using campus-based financial aid strategically. In D. Hossler, & B. Bontrager (Eds.), *Handbook of Strategic Enrollment Management* (pp. 213-227). 213-227.

Buckley, J., Letukas, L., & Wildavsky, B. (Eds.). (2018). Measuring success: Testing, grades, and the future of college admissions. Johns Hopkins University Press.

Bulmer, M. G. (1979). Principles of statistics. Courier Corporation.

Burton, N., & Ramist, L. (2001). Predicting success in college: SAT studies of classes graduating since 1980. (College Board 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 Note No. RN-10). New York: College Board.

Camara, W. J., & Kimmel, E. W. (2005). Choosing students: Higher education admissions tools for the 21st Century. Mahwah, N.J.: Lawrence Erlbaum Associates, Inc.

Camara, W. J., & Kimmel, E., Scheuneman, J., & Sawtell, E. (2003). *Whose grades are inflated?* (College Board Research Report. No. 03-04). New York: College Board.

Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt, & W. Borman (Eds.), *Personnel Selection in Organizations* (pp. 35-70). San Francisco: Jossey-Bass.

Clinedinst, M. E., & Hawkins, D. A. (2009). State of college admission. Washington, DC: *National Association for College Admission Counseling*.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37-46.

Compas, B., Connor-Smith, J., Saltzman, H., Thomsen, A., & Wadsworth, M. (2001). Coping with stress during childhood and adolescence: Problems, progress, and potential in theory and research. *Psychological Bulletin*, *127*(1), 87-127.

Condition of College and Career Readiness (2016). ACT. Retrieved from http://www.act.org/content/dam/act/unsecured/documents/CCCR_National_2016.pdf.

Costa, P. T., & McCrae, R. R. (1992). NEO PI-R: Professional manual: Revised NEO PI-R and NEO-FFI. Florida: Psychological Assessment Resources, Inc.

Costa, P. T. & McCrae, R. R. (1994). Stability and change in personality from adolescence through adulthood. In C. F. Halverson, Jr., G. A. Kohnmstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 139-155). Mahwah, NJ: Lawrence Erlbaum Associates.

Covington, M. (2000). Goal theory, motivation, and school achievement: An integrative review. *Annual Review of Psychology*, *51*, 171-200.

Davenport, Z., Martinez-Saenz, M., & Rhine, L. (2012). The student success conundrum. In B. Bontrager, D. Ingersoll, & R. Ingersoll (Eds.), *Strategic Enrollment Management: Transforming Higher Education* (pp. 25-50). Washington, DC: American Association of Collegiate Registrars and Admissions Officers.

Digman, J. (1997). Higher order factors of the Big Five. *Journal of Personality and Social Psychology*, *73*, 1246-1256.

Dikli, S. (2006). An overview of automated scoring of essays. *The Journal of Technology, Learning and Assessment*, *5*(1).

DiPerna, J. C., & Elliott, S. N. (1999). Development and validation of the academic competence evaluation scales. *Journal of Psychoeducational Assessment*, *17*(3), 207-225.

Drasgow, F., & Kang, T. (1984). Statistical power of differential validity and differential prediction analyses for detecting measurement nonequivalence. *Journal of Applied Psychology*, *69*, 498–508.

Dweck, C. S., Chiu, C, & Hong, Y. (1995). Implicit theories and their role in judgments and reactions: A world from two perspectives. *Psychological Inquiry*, *6*, 267-285.

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

Expected Family Contribution (EFC). (2017). U.S. Department of Education. Retrieved from https://fafsa.ed.gov.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861-874.

Fisher v. University of Texas at Austin, 133 S. Ct. 2411, 570 U.S., 186 L. Ed. 2d 474 (2013).

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science*, *15*(6), 373-378.

Gardner, H. (2011). Frames of mind: The theory of multiple intelligences (3rd ed.). New York: Basic Books.

Geiser, S., & Santelices, M. (2007). Validity of high-school grades in predicting student success beyond the freshman year: High-school record vs. standardized tests as indicators of four-year college outcomes. *IDEAS Working Paper Series from RePEc*, IDEAS Working Paper Series from RePEc, 2007.

Geiser, S. & Studley, R. (2002). UC and the SAT: Predictive validity and differential impact of the SAT I and SAT II at the University of California. *Educational Assessment*, *8*(1): 1-26.

Gratz v. Bollinger, 539 U.S. 244, 123 S. Ct. 2411, 156 L. Ed. 2d 257 (2003).

Grutter v. Bollinger, 539 U.S. 306, 123 S. Ct. 2325, 156 L. Ed. 2d 304 (2003).

Hagedorn, L. S. (2005). How to define retention: A new look at an old problem. In A. Seidman (Ed.), *College Student Retention: Formula for Student Success* (pp. 89-105). Westport, CT: Praeger Publishers.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23.

Haracklewicz, J., Barron, K., Tauer, J., Elliot, A., & Pressley, G. Michael. (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.

Haveman, R., & Smeeding, T. (2006). The role of higher education in social mobility. *The Future of Children*, *16*(2), 125-150. doi:10.1353/foc.2006.0015.

Hezlett, S. A., Kuncel, N., Vey, M. A., Ahart, A. M. Ones, D. S., Campbell, J. P. & Camara, W. J. (2001, April) *The effectiveness of the SAT in predicting success early and late in college: A comprehensive meta-analysis*. Paper presented at the National Council of Measurement in Education, Seattle, WA.

Hiss, W., & Franks, W. (2014). Defining promise: Optional standardized testing policies in American college and university admissions. Report of the National Association for College Admission Counseling (NACAC). Retrieved from http://www.nacacnet.org/research/research-data/nacacresearch/Documents/DefiningPromise.pdf.

Hong, Y., Chiu, C., Lin, D., Wan, W., Dweck, C., & Diener, Ed. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology*, *77*(3), 588-599.

Hossler, D., & Bontrager, B. (2014). Handbook of strategic enrollment management. Hoboken: Wiley.

Judge, T. A., Jackson, C. L., Shaw, J. C., Scott, B. A., & Rich, B. L. (2007). Self-efficacy and work-related performance: The integral role of individual differences. *Journal of Applied Psychology*, *92*, 107−127.

Kaplan, A., Pendergast, L., French, B., Kanno, Y. (2017). Development of a measure of test-optional applicants' motivational-developmental attributes. Institute of Educational Sciences Proposal.

Kegan, R. (1994). In over our heads: The mental demands of modern life. Cambridge, Mass.: Harvard University Press.

Komarraju, M., Karau, S. J., & Schmeck, R. R. (2009). Role of the big five personality traits in predicting college students' academic motivation and achievement. *Learning and Individual Differences*, *19*(1), 47-52. doi:10.1016/j.lindif.2008.07.001.

Kobrin, J. L., & Patterson, B. F. (2011). Contextual factors associated with the validity of SAT scores and high school GPA for predicting first-year college grades. *Educational Assessment*, *16*(4), 207-226.

Kobrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). Validity of the SAT for predicting first-year college grade point average (College Board Research Report No. 2008-5). New York: College Board.

Kobrin, J. L., Sathy, V., & Shaw, E. J. (2006). A historical view of subgroup performance differences on the SAT (College Board Research Report No. 2006-5). New York: College Board.

Le, H., Casillas, A., Robbins, S., & Langley, R. (2005). Motivational and skills, social, and self-management predictors of college outcomes: Constructing the student readiness inventory. *Educational and Psychological Measurement*, *65*(3), 482-508. doi:10.1177/0031364404272493.

Lens, W. (1986). Future time perspective: A cognitive-motivational concept. In Brown, D. R., and Veroff, J. (Eds.), *Frontiers of Motivational Psychology*, Springer, New York, pp. 173–190.

Liebetrau, A. M. (1983). Quantitative applications in the social sciences: Measures of association Thousand Oaks, CA: Sage Publications.

Light, R. J. (1971). Measures of response agreement for qualitative data: Some generalizations and alternatives. Psychological Bulletin, 76(5), 365-377.

Linn, R. L. (1978). Single-group validity, differential validity, and differential prediction. *Journal of Applied Psychology*, *63*, 507–512.

Linn, R. L. (2005). Evaluating college applicants: Some alternatives. In Choosing students: Higher education admission tools for the 21st Century. Mahwah, NJ: Lawrence Erlbaum Associates, 109-25.

Little, R., & Rubin, D. (2014). Statistical analysis with missing data (2nd ed., Wiley Series in Probability and Statistics). Hoboken: Wiley.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57*(9), 705.

Lucido, J. A. (2014). How admission decisions get made. In Hossler, D. & Bontrager, B. (Eds.), *Handbook of strategic enrollment management* (pp. 147-173). San Francisco: Jossey-Bass.

Maller, S. J., & Pei, L. K. (2017). Best practices in detecting bias in cognitive tests. In McCallum, R. (Ed.), *Handbook of nonverbal assessment* (pp. 21-37). Springer, Cham.

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, *81*(1), 59-77.

Mattern, K. D., & Patterson, B. F. (2013). Test of slope and intercept bias in college admissions: A response to Aguinis, Culpepper, and Pierce (2010). *Journal of Applied Psychology*, *98*, 134–147.

McCluckie, B. (2014). Identifying students 'at risk' of withdrawal using ROC analysis of attendance data. *Journal of Further and Higher Education*, *38*(4), 523-535.

McHugh, M. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276-282.

Mertens, D. M. (2014). Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods. Thousand Oaks, CA: Sage Publications.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: Macmillan.

Milem, J. F., & Berger, J. B. (1997). A modified model of college student persistence: Exploring the relationship between Astin's theory of involvement and Tinto's theory of departure. *Journal of College Student Development*, *38*, 387-400.

Morgan, R. (1989). Analysis of the predictive validity of the SAT and high school grades from 1976 to 1983 (College Board Report. No. 89–7). New York: College Board.

Mulugetta, Yuko. (2013). Going test-optional: Gathering evidence and making the decision at Ithaca College. http://www.ithaca.edu/ir/docs/testoptionalpaper.pdf.

National Association for College Admissions Counseling. (2016). Use of predictive validity studies to inform admission practices. http://www.nacacnet.org/research/research-data/nacac-research/Documents/TestValidity.pdf.

Neville, H. A., Heppner, P., & Wang, L. (1997). Relations among racial identity attitudes, perceived stressors, and coping styles in African American college students. *Journal of Counseling and Development*, *75*, 303-311.

Nisbet, J., Ruble, V. E., & Shurr, K. T. (1982). Predictors of academic success with high risk college students. *Journal of College Student Personnel*, *23*, 227-235.

Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology*, *93*(1), 116.

Pascarella, E., & Terenzini, P. (1991). How college affects students: Findings and insights from twenty years of research. San Francisco, CA: Jossey-Bass.

Pell Grant. (2018). U.S. Department of Education. Retrieved from https://www.ed.gov/programs/fpg

Pendergast, L. L., von der Embse, N., Kilgus, S. P., & Eklund, K. R. (2017). Measurement equivalence: A non-technical primer on categorical multi-group confirmatory factor analysis in School Psychology. *Journal of School Psychology*, *60*, 65-82. doi: 10.1016/j.jsp.2016.11.002.

Pearson, K. (1895). Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *186*(0), 343-414.

Perfetto, G., Escandón, M., Graff, S., Rigol, G., & Schmidt, A. (1999). Toward a taxonomy of the admissions decision-making process: A public document based on the first and second College Board conferences on admissions models. College Entrance Examination Board.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*(1), 33.

Rohde, T. E., & Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence*, *35*(1), 83-92.

Robbins, S. B., Allen, J., Cassillas, A., Peterson, C. H., & Le, H. (2006). Unraveling the differential effects of motivational and skills, social and self-management measures from traditional predictors of college outcomes. *Journal of Educational Psychology*, *98*, 598-616.

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

Rothstein, J. (2004). College performance predictions and the SAT. *Journal of Econometrics*, *121*, 297–317.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68-78.

Sacket, P. R. (2005). The performance-diversity tradeoff in admissions testing. In *Choosing students: Higher education admission tools for the 21st Century*. Mahwah, NJ: Lawrence Erlbaum Associates, 109-25.

Sackett, P. R., Kuncel, N., Arneson, J., Cooper, S., & Waters, S. (2009). Does socioeconomic status explain the relationship between admissions tests and post-secondary academic performance? *Psychological Bulletin*, *135*, 1-22.

Sackett, P., Kuncel, N., Beatty, A., Rigdon, J., Shen, W., & Kiger, T. (2012). The role of socioeconomic status in SAT-grade relationships and in college admissions decisions. *Psychological Science*, *23*(9), 1000-1007.

Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. (2001). High-stakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative-action world. *American Psychologist*, *56*, 302-318.

Salkind, N. J. (2007). Encyclopedia of measurement and statistics (Vol. 1). Thousand Oaks, CA: Sage Publications.

SAT 2016 College-bound seniors total group profile report. (2016). Retrieved from https://reports.collegeboard.org/pdf/total-group-2016.pdf.

Sedlacek, W. E. (1986). Sources of method bias in test bias research. In *Measures in the admissions process* (pp. 86-92). New York: College Entrance Examination Board.

Sedlacek, W. E. (2004). Beyond the big test: Noncognitive assessment in higher education. San Francisco: Jossey-Bass.

Shanley, B. J. (2007). Test-optional admission at a liberal arts college: A founding mission affirmed. *Harvard Educational Review*, *77*(4), 429.

Schunk, D. H. (1981). Modeling and attributional effects on children's achievement: A self-efficacy analysis. *Journal of Educational Psychology*, *73*, 93–105.

Silvia, P. J., & Duval, T. S. (2001). Objective self-awareness theory: Recent progress and enduring problems. *Personality and Social Psychology Review*, *5*(3), 230-241.

Sink, C. A., & Mvududu, N. H. (2010). Statistical power, sampling, and effect sizes: Three keys to research relevancy. *Counseling Outcome Research and Evaluation*, *1*(2), 1-18.

Soares, J. A. (2012). SAT Wars: The case for test-optional college admissions. New York, NY: Teachers College Press.

Spence, J. T., & Helmreich, R. L. (1983). Achievement-related motives and behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives: Psychological and sociological approaches* (pp. 7–74). San Francisco: Freeman.

Sternberg, R. J. (1985). Beyond IQ: A triarchic theory of human intelligence. New York: Cambridge University Press.

Sternberg, R. J. (1997). Successful intelligence. New York, NY: Plume.

Sternberg, R. J. (2010). College admissions for the 21st Century. Cambridge, MA: Harvard University Press.

Swets, J. A., Dawes, R. M., & Monahan, J. (2000). Psychological science can improve diagnostic decisions. *Psychological Science in the Public Interest*, *1*(1), 1-26.

Syverson, S. (2007). The role of standardized tests in college admissions: Test-Optional admissions. *New Directions for Student Services*, 2007(118), 55-70. doi:10.1002/ss.241.

Thomas, L. L., Kuncel, N. R., & Credé, M. (2007). Noncognitive variables in college admissions. *Educational and Psychological Measurement*, *67*(4), 635.

Ting, S. R., & Robinson, T. L. (1998). First-year academic success: A prediction combining cognitive and psychosocial variables for White and African American students. *Journal of College Student Development*, *39*, 599-610.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, *45*, 89-125.

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

Tracey, T. J., & Sedlacek, W. E. (1984). Noncognitive variables in predicting academic success by race. *Measurement & Evaluation in Guidance*, *16*, 171-178.

Tracey, T. J., & Sedlacek, W. E. (1985). The relationship of noncognitive variables to academic success: A longitudinal comparison by race. *Journal of College Student Personnel*, *26*, 405-410.

Trapmann, S., Hell, B., Hirk, J. W., & Schuler, H. (2007). Meta-analysis of the relationship between the Big Five and academic success at university. *Zeitschrift Fur Psychologie*, 215, 132–151.

Tross, S. A., Harper, J. P., Osher, L. W., & Kneidinger, L. M. (2000). Not just the usual cast of characteristics: Using personality to predict college performance and retention. *Journal of College Student Development*, *41*(3), 323-334.

University of California Regents v. Bakke, 438 U.S. 265, 98 S. Ct. 2733, 57 L. Ed. 2d 750 (1978).

U.S. Department of Education (2017). Institute of Education Sciences, National Center for Education Statistics.

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

Vansteenkiste, M., & Ryan, R. M. (2013). On psychological growth and vulnerability: Basic psychological need satisfaction and need frustration as a unifying principle. *Journal of Psychotherapy Integration*, *23*, 263–280. doi: 10.1037/a0032359

Vars, F. E., & Bowen, W. G., (1998). Scholastic Aptitude Test scores, race, and academic performance in selective colleges and universities. In C. Jencks & M. Phillips (Eds.), The Black-White test score gap (pp. 457-459). Washington, DC: Brookings Institution Press.

Walpole, M. (2003). Socioeconomic Status and College: How SES Affects College Experiences and Outcomes. *The Review of Higher Education*, *27*(1), 45-73. The Johns Hopkins University Press.

Weiner, B. (2010). The development of an attribution-based theory of motivation: A history of ideas. *Educational Psychologist*, *45*(1), 28-36. doi:10.1080/00461520903433596.

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

Willingham, W. W. (1985). Success in college: The role of personal qualities and academic abilities. New York: College Board.

Willingham, W. W. Lewis, C., Morgan, R. & Ramist, L. (1990), Predicting college grades: An analysis of trends over two decades. Princeton, NJ: Educational Testing Service, 253–288.

Wilson, K. M. (1983). A review of research on the prediction of academic performance after the freshman year (College Board Report No. 83-2). New York: College Board.

Wolfe, R. N., & Johnson, S. D. (1995). Personality as a predictor of college performance. *Educational and Psychological Measurement*, *55*(2), 177-185. doi:10.1177/0013164495055002002.

Yap, B. W., & Sim, C. H. (2011). Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation*, *81*(12), 2141-2155.

Young, J. W., & Kobrin, J. L. (2001). Differential validity, differential prediction, and college admission testing: A comprehensive review and analysis (College Board Research Report No. 2001-6). New York: College Board.

Youngstrom, E. A. (2013). A primer on receiver operating characteristic analysis and diagnostic efficiency statistics for pediatric psychology: We are ready to ROC. *Journal of Pediatric Psychology*, *39*(2), 204-221.

Yu, C. H., DiGangi, S., Jannasch-Pennell, A., & Kaprolet, C. (2010). A data mining approach for identifying predictors of student retention from sophomore to junior year. *Journal of Data Science*, *8*(2), 307-325.

Zimmerman, B. J. (1989). A social-cognitive view of self-regulated academic learning. *Journal of Educational Psychology*, *81*, 329–339.

Zwick, R. (2002). Fair game? The use of standardized admissions tests in higher education. New York: New York.

APPENDIX A

NEW STUDENT QUESTIONNAIRE (NSQ)

Note. "City University" was used to ensure the anonymity of the participating institution.

1. Is English your native language?

    a. Yes

    b. No

2. During the school year how many hours do you plan to work (for money) per week?

    a. None

    b. 1 to 15 hours

    c. 16 to 20 hours

    d. 21 to 25 hours

    e. More than 25 hours

3. In what kind of residence will you be living during your first semester at City University?

    a. University-owned housing (including residence halls)

    b. Home of parents/relatives

    c. Your own home or apartment

    d. With other family members

    e. Other

4. Where is your permanent home?

    a. City of Philadelphia

    b. Suburban Philadelphia

    c. Pennsylvania, other than suburban Philadelphia

    d. United States, other than Pennsylvania

    e. Other country

5. What is your U.S. Military status?

a. No military service

b. Active military service

c. Veteran

d. Reserves or ROTC

e. Other

6. What is your best estimate of the total income of your PARENTAL FAMILY during the past year?

a. Less than $20,000

b. $20,000 to $39,999

c. $40,000 to $59,000

d. $60,000 to $79,000

e. $80,000 to $99,000

f. $100,000 to $124,000

g. $125,000 to $149,000

h. $150,000+

7. What is your best estimate of YOUR OWN total income during the past year?

a. Less than $2,000

b. $2,000 to $3,999

c. $4,000 to $5,999

d. $6,000 to $7,999

e. $8,000 or more

8. Did you apply for financial aid from City University for this year?

a. Yes, and I received an aid package

b. Yes, but I did not receive aid

c. Yes, but I have yet to hear about an aid package

d. No, but I intend to apply

e. No, I will not need financial aid to attend City University

9. How did the amount of financial aid you received from City University compare to other schools to which you were admitted?

a. Higher than most

b. About the same

c. Lower than most

d. Applied for aid only at the institution

e. Did not apply for financial aid

10. Do you have any concern about your ability to finance your college education?

a. None (I am confident that I will have sufficient funds)

b. Some concern (but I will probably have enough funds)

c. Major concern (not sure I will have enough funds to complete college)

11. What was your rating of City University at the time you applied for admission?

a. The institution was my first choice

b. The institution was my second choice

c. The institution was my third or lower choice

12. What is the highest level of formal education completed by your father or guardian 1?

a. Did not graduate high school

b. Graduated from high school

c. Some college education

d. Graduated from college (a bachelor's degree)

e. Postgraduate or professional degree

13. What is the highest level of formal education completed by your mother or guardian 2?

a. Did not graduate high school

b. Graduated from high school

c. Some college education

d. Graduated from college (a bachelor's degree)

d. Postgraduate or professional degree

14. What is the highest level of formal education completed by any of your college age (18 years of age or older) siblings?

a. Did not graduate from high school

b. Graduated from high school

c. Some college education, but did not graduate

d. Currently enrolled in college

e. Graduated from college (a bachelor's degree or higher)

f. I do not have any college age siblings

15. What was your approximate high school average?

a. A

b. B+

c. B

d. B-

e. C+ or lower

16. What scholastic average do you expect to obtain in college?

    a. A

    b. B+

    c. B

    d. B-

    e. C+ or lower

17. Scholastically, where did you rank in your high school graduating class?

    a. Top 10%

    b. Top 20%

    c. Top 30%

    d. Top 50%

    e. Not among top 50%

18. In general, how well do you feel that your high school prepared you to do college work?

    a. Very well

    b. Fairly well

    c. Uncertain

    d. Poorly

    e. Very poorly

19. Have you decided on an academic major?

    a. Yes

    b. No

20. Do you consider yourself a person who has a disability?

    a. Yes

b. No

c. Uncertain

21. During the school year, on the average, how many hours do you plan to study per week?

a. None

b. 1 to 15 hours

c. 16 to 20 hours

d. 21 to 25 hours

e. More than 25 hours

Questions 22-25: During high school (grades 9-12), how many years did you study each of the following subjects?

For questions 22 through 25 use the following responses:

a. None

b. One

c. Two

d. Three

e. Four

22. English

23. Mathematics

24. Foreign Language

25. Natural Sciences

Questions 26-29: During high school (grades 9-12), on average, what was your grade in the following subjects?

For questions 26 through 29 use the following responses:

a. A

b. B+

c. B

d. B-

e. C+ or lower

26. English

27. Mathematics

28. Foreign Languages

29. Natural Sciences

Questions 30-36: How important were the following in your decision to go to college?

For questions 30 through 36 use the following responses:

a. Very important

b. Somewhat important

c. Not important

30. I wanted to get a general education.

31. My family wanted me to go.

32. I wanted to get away from home.

33. I wanted to be able to get a better job.

34. I wanted to learn more about things that interest me.

35. I wanted to prepare myself for graduate or professional school.

36. It seemed like a good thing to do as a transition to work.

Questions 37-44: How important were the following in your finding out about or selecting City University?

For questions 37 through 44 use the following responses:

a. Very important

b. Somewhat important

c. Not important

d. Does not apply/did not attend

37. Personal call from City University student

38. City University social media channels (Facebook, Twitter, Instagram, etc.)

39. City University brochure or mailings

40. E-mail communication from City University

41. City University's website

42. City University open house or reception

43. Regular campus visit

44. High school visit by City University representative or college fair

Questions 45-58: Below are some reasons that might have influenced your decision to attend City University. How important was each reason in your decision to come here?

For questions 45 through 58 use the following responses:

a. Very important positive factor

b. Somewhat important positive factor

c. Not a positive factor

45. Affordable tuition

46. City University's student body size

47. Social atmosphere

48. Closeness to home

49. Urban location

50. Variety of academic programs available

51. Variety of co-curricular programs available

52. Reputation of City University

53. Reputation of your specific major at City University

54. Advice and experience of family

55. Advice and experience of friends

56. Meeting students with backgrounds and interests similar to yours

57. Meeting students with backgrounds and interests different from yours

58. Availability of financial aid

Questions 59-73: What is the chance that you will do the following while you are at City University?

For questions 59 through 73 use the following responses:

a. Very good chance

b. Some chance

c. Very little chance

d. No chance

59. Change your major field of study

60. Be a student leader

61. Work full time while attending college

62. Join a social organization or club

63. Play varsity / intercollegiate athletics

64. Need more than 4 years to complete degree requirements

65. Make close friends

66. Work with a professor on a research project

67. Receive encouragement from family while you're in college

68. Get tutoring help in specific courses

69. Transfer to another college before graduating

70. Take an on-line course

71. Find a job after college in your chosen field of study

72. Participate in volunteer or community service work

73. Study abroad/away

Questions 74-82: Please indicate your level of agreement with each of the following statements:

For questions 74 through 82 use the following responses:

a. Definitely agree

b. Somewhat agree

c. Neither agree nor disagree

d. Somewhat disagree

e. Definitely disagree

74. I want to live and work in Philadelphia after graduation.

75. Most of my teachers considered me one of the harder workers in their class.

76. I find it difficult to keep a plan of action in my school work.

77. I enjoy studying and reading about things on which I am working.

78. I know how to manage my time well.

79. I am self-confident.

80. My plans have frequently seemed so full of difficulties that I have had to give them up.

81. I am organized and have good study habits.

82. I prefer to be independent of others in deciding what I want to do.

APPENDIX B

NSQ MOTIVATIONAL ITEMS

New Student Questionnaire (NSQ) Item	Abbreviation
75. Most of my teachers considered me one of the harder workers in their class.	Hard Work
76. I find it difficult to keep a plan of action in my school work.	Difficult Plan
77. I enjoy studying and reading about things on which I am working.	Enjoy Study
78. I know how to manage my time well.	Time Management
79. I am self-confident.	Confident
80. My plans have frequently seemed so full of difficulties that I have had to give them up.	Give Up
81. I am organized and have good study habits.	Organized
82. I prefer to be independent of others in deciding what I want to do.	Independent

APPENDIX C

DEMOGRAPHIC REPRESENTITIVENESS OF STUDY PARTICIPANTS

	Total Test-Optional Study Participants N = 1,610		2015-2016 National Undergraduate Postsecondary Participants N = 10,988,573	
Demographic Variable	Frequency	Percentage	Frequency	Percentage
Gender				
Female	1,115	69.3	6,080,885	55.3
Male	495	30.7	4,907,688	44.7
Race				
African American	359	22.3	1,364,542	12.4
Asian	190	11.8	657,507	6.0
Hispanic	158	9.8	1,370,383	12.5
International	14	.90	509,437	4.6
White	757	47.0	5,978,092	54.4
Other	132	8.2	1,108,612	10.1

Note. National data includes undergraduate enrollment demographics of 1,964 degree-granting institutions categorized by the National Center for Education Statistics Integrated Postsecondary Data System (2017) as "degree-granting, primarily baccalaureate or above" located in the United States.

APPENDIX D

INTENDED ACADEMIC DISCIPLINE BY APPLICATION TYPE

Intended Academic Discipline for Study 1 Participants

	Test-Optional Applicants		Standard Applicants	
Academic Discipline	Frequency	Percentage	Frequency	Percentage
Art	42	4.7	184	4.6
Business	158	17.8	895	22.2
Communications	75	8.5	284	7.0
Education	37	4.2	74	1.8
Engineering	25	2.8	351	8.7
Liberal Arts	135	15.2	540	13.4
Performing Arts	49	5.5	217	5.4
Public Health	119	13.4	401	9.9
Science	176	19.9	849	21.0
Undeclared/Undecided	70	7.9	239	5.9

Note. Business includes Sport, Tourism, and Hospitality. Liberal Arts includes Environmental Design. Performing Arts includes Music, Theater, Film, and Media Arts. Public Health includes Social Work.

Intended Academic Discipline for Study 2 Participants

Academic Discipline	Test-Optional Applicants		Standard Applicants	
	Frequency	Percentage	Frequency	Percentage
Art	43	5.9	216	4.8
Business	130	18.0	1,030	22.9
Communications	63	8.7	297	6.6
Education	25	3.5	88	2.0
Engineering	24	3.3	357	7.9
Liberal Arts	111	15.3	623	13.8
Performing Arts	53	7.3	314	7.0
Public Health	99	13.7	396	8.8
Science	121	16.7	948	21.0
Undeclared/Undecided	55	7.6	236	5.2

Note. Business includes Sport, Tourism, and Hospitality. Performing Arts includes Music, Theater, Film, and Media Arts. Public Health includes Social Work.

Test-Optional Applicants as a Percentage of Total Applicants by Academic Discipline

Academic Discipline	Study 1	Study 2
Art	18.6	16.6
Business	15.0	11.2
Communications	21.0	17.5
Education	67.6	22.1
Engineering	6.6	6.3
Liberal Arts	20.0	15.1
Performing Arts	18.4	14.4
Public Health	22.9	20.0
Science	17.2	11.3
Undeclared/Undecided	22.7	18.9

Note. Business includes Sport, Tourism, and Hospitality. Liberal Arts includes Environmental Design. Performing Arts includes Music, Theater, Film, and Media Arts. Public Health includes Social Work.

APPENDIX E

DESCRIPTIVE STATISTICS OF PREDICTOR VARIABLES

Descriptive Statistics of Predictor Variables in Study 1

Predictor Variable	N	Range	Mean	Standard Deviation	Skewness		Kurtosis	
					Statistic	Standard Error	Statistic	Standard Error
Test-Optional Essay Rating Score	889	18 – 33	25.70	2.60	-.11	.08	.01	.16
SAT Total	4,342	530 – 1600	1,145.15	157.59	.05	.04	.02	.07
SAT Verbal	3,972	220 – 800	562.00	85.95	.01	.04	.17	.08
SAT Quantitative	3,972	230 – 800	573.30	87.52	.08	.04	.11	.08
SAT Writing	3,970	250 – 800	553.31	89.73	.07	.04	-.09	.08
High School Grade Point Average	4,853	1.70 – 4.00	3.51	.36	-.66	.04	.32	.07

Descriptive Statistics of Predictor Variables in Study 2

Predictor Variable	N	Range	Mean	Standard Deviation	Skewness		Kurtosis	
					Statistic	Standard Error	Statistic	Standard Error
Test-Optional Essay Rating Score	734	13 – 32	23.23	3.47	.05	.09	-.31	.18
SAT Total	4,674	590 – 1600	1,168.18	161.68	-.05	.04	-.08	.07
SAT Verbal	4,165	280 – 800	570.72	86.76	.03	.04	-.05	.08
SAT Quantitative	4,165	230 – 800	582.87	89.44	.04	.04	-.04	.08
SAT Writing	4,163	230 – 800	562.00	88.71	.10	.04	-.08	.08
High School Grade Point Average	5,162	2.00 – 4.00	3.56	.33	-.71	.03	.26	.07

APPENDIX F

DESCRIPTIVE STATISTICS OF OUTCOME VARIABLES

Descriptive Statistics of Outcome Variables in Study 1

Predictor Variable	N	Range	Mean	Standard Deviation	Skewness		Kurtosis	
					Statistic	Standard Error	Statistic	Standard Error
Cumulative Undergraduate Grade Point Average	4,919	.00 – 4.00	3.07	.73	-1.65	.04	3.74	.07
Semesters Retained	4,920	0 – 4	3.72	.76	-2.67	.04	5.93	.07

Descriptive Statistics of Outcome Variables in Study 2

Predictor Variable	N	Range	Mean	Standard Deviation	Skewness		Kurtosis	
					Statistic	Standard Error	Statistic	Standard Error
Cumulative Undergraduate Grade Point Average	5,229	.00 – 4.00	3.14	.74	-1.69	.03	3.78	.07
Semesters Retained	5,229	0 – 2	1.96	.21	-4.54	.03	19.43	.07

APPENDIX G

COORDINATES OF THE ROC CURVE FOR RETENTION

Coordinates of the ROC Curve[a] for Retention in Study 1

Positive if Greater Than or Equal To[b]	Sensitivity	1 - Specificity
16.880	1.000	1.000
18.065	.999	1.000
18.440	.997	1.000
18.815	.992	1.000
19.190	.988	1.000
19.565	.987	.993
19.940	.984	.972
20.315	.979	.957
20.690	.971	.929
21.065	.961	.908
21.440	.949	.879
21.815	.941	.865
22.190	.926	.858
22.565	.902	.830
22.940	.882	.794
23.315	.842	.773
23.690	.794	.716
24.065	.758	.688
24.440	.707	.645
24.815	.660	.610
25.190	.586	.560
25.565	.529	.489
25.940	.463	.397
26.315	.420	.355
26.690	.361	.291

27.065	.311	.262
27.440	.262	.220
27.815	.211	.206
28.190	.166	.170
28.565	.136	.135
28.940	.106	.106
29.315	.083	.078
29.690	.060	.043
30.065	.047	.028
30.440	.028	.028
30.815	.023	.021
31.190	.013	.014
31.565	.012	.007
31.940	.008	.000
32.315	.007	.000
32.690	.004	.000
33.880	.000	.000

Note. a. The test result variable(s): The test-optional essay rating score has at least one tie between the positive actual state group and the negative actual state group.
b. The smallest cutoff value is the minimum observed test value minus 1, and the largest cutoff value is the maximum observed test value plus 1. All other cutoff values are the averages of two consecutive ordered observed test values.

Coordinates of the ROC Curve for Retention in Study 2

Positive if Greater Than or Equal To[b]	Sensitivity	1 - Specificity
12.630	1.000	1.000
13.815	.999	1.000
14.125	.997	1.000
14.315	.996	.978
14.440	.994	.978
14.690	.993	.978
15.005	.991	.978
15.380	.990	.978
15.815	.987	.978
16.250	.985	.978
16.565	.984	.978
16.690	.981	.978
16.815	.978	.978
16.940	.975	.978
17.065	.974	.957
17.190	.969	.957
17.315	.968	.935
17.440	.965	.913
17.565	.958	.913
17.690	.956	.913
17.815	.951	.913
17.940	.945	.913
18.065	.942	.891
18.190	.938	.891
18.315	.926	.870
18.440	.920	.870
18.565	.901	.870
18.690	.890	.870

18.815	.887	.848
18.940	.882	.826
19.065	.879	.826
19.190	.876	.826
19.315	.863	.783
19.440	.855	.783
19.565	.849	.739
19.690	.836	.739
19.815	.831	.739
19.940	.827	.739
20.065	.814	.696
20.190	.807	.696
20.315	.798	.696
20.440	.786	.652
20.565	.783	.652
20.690	.772	.630
20.815	.754	.565
20.940	.746	.565
21.065	.738	.565
21.190	.727	.565
21.315	.706	.565
21.440	.698	.565
21.565	.672	.543
21.690	.660	.522
21.815	.654	.522
21.940	.635	.478
22.065	.629	.478
22.190	.621	.478
22.315	.597	.413
22.440	.593	.413
22.565	.577	.391

22.690	.570	.391
22.815	.563	.391
22.940	.554	.391
23.065	.528	.370
23.190	.522	.348
23.315	.500	.348
23.440	.490	.326
23.565	.480	.326
23.690	.467	.326
23.815	.432	.283
23.940	.427	.283
24.065	.410	.283
24.190	.400	.283
24.315	.384	.283
24.440	.372	.261
24.565	.349	.196
24.690	.343	.196
24.815	.336	.152
24.940	.324	.152
25.065	.308	.109
25.190	.302	.087
25.315	.282	.065
25.440	.275	.065
25.565	.265	.065
25.690	.260	.065
25.815	.240	.065
25.940	.233	.065
26.065	.222	.065
26.190	.214	.065
26.315	.208	.065
26.440	.199	.065

26.565	.188	.065
26.690	.183	.065
26.815	.163	.065
26.940	.158	.065
27.065	.150	.065
27.190	.140	.065
27.375	.126	.065
27.565	.113	.065
27.690	.112	.043
27.815	.105	.022
27.940	.097	.022
28.065	.086	.022
28.190	.084	.022
28.375	.073	.022
28.625	.067	.022
28.815	.058	.022
28.940	.054	.022
29.065	.052	.022
29.190	.051	.022
29.315	.042	.022
29.440	.041	.022
29.565	.038	.022
29.690	.033	.022
29.875	.031	.022
30.065	.029	.022
30.255	.026	.022
30.440	.025	.000
30.625	.019	.000
30.815	.017	.000
30.940	.016	.000
31.065	.012	.000

31.315	.010	.000
31.565	.009	.000
31.690	.007	.000
31.815	.006	.000
31.940	.004	.000
32.065	.003	.000
32.190	.001	.000
33.250	.000	.000

Note. a. The test result variable(s): The test-optional essay rating score has at least one tie between the positive actual state group and the negative actual state group.
b. The smallest cutoff value is the minimum observed test value minus 1, and the largest cutoff value is the maximum observed test value plus 1. All other cutoff values are the averages of two consecutive ordered observed test values.