

St. John's University

St. John's Scholar

Theses and Dissertations

2021

THE PSYCHOMETRIC PROPERTIES OF A SOCIAL-EMOTIONAL LEARNING MEASURE

Allison Murray

Follow this and additional works at: https://scholar.stjohns.edu/theses_dissertations



Part of the Psychology Commons

THE PSYCHOMETRIC PROPERTIES OF A SOCIAL-EMOTIONAL LEARNING
MEASURE

A dissertation submitted in partial fulfillment

of the requirements for the degree of

DOCTOR OF PSYCHOLOGY

to the faculty of the

DEPARTMENT OF PSYCHOLOGY

of

ST. JOHN'S COLLEGE OF LIBERAL ARTS AND SCIENCES

at

ST. JOHN'S UNIVERSITY

New York

by

Allison M. Murray

Date Submitted: _____

Date Approved: _____

Allison M. Murray

Marlene Sotelo-Dynega

© Copyright by Allison Murray 2021
All Rights Reserved

ABSTRACT

THE PSYCHOMETRIC PROPERTIES OF A SOCIAL-EMOTIONAL LEARNING MEASURE

Allison M. Murray

Each year many students take college admissions exams (i.e., SAT® and ACT®), hoping to demonstrate their ability to perform at a collegiate level and gain admission to desired universities. However, a growing movement encourages colleges and universities to abandon this practice in their admissions protocol and instead consider alternative factors, such as, social-emotional learning skills, to identify promising applicants. As such, this study examined the psychometric properties of a novel social-emotional learning measure, ACT® Tessera®, which conceptualizes social-emotional traits through the Five-Factor Model lens using different measurement methods (Self Report Likert, Situational Judgement Tests, Forced Choice). Using data obtained from an undergraduate student sample at a metropolitan university, reliability and validity analyses revealed promising evidence for the scale's ability to measure social-emotional skills. However, recommendations for future scale iterations are made to improve the scales' psychometric properties. Then, ACT® Tessera® social-emotional trait measures were assessed alongside traditional college achievement predictors (intelligence, cognitive ability, standardized test scores) to determine their ability to predict undergraduate success. Preliminary evidence provided by this study suggests that considering social-emotional traits in conjunction with high school GPA may provide useful predictions of university

success, without standardized test scores. Suggestions for future research and implications for school psychologists are discussed.

ACKNOWLEDGEMENTS

There are not enough words to express how truly grateful I am for each person who supported my journey and helped me to get to this point in my graduate career.

First and foremost, I would like to thank my dissertation mentor, Dr. Marlene-Sotelo-Dynega. From the day I started the program, you were a constant source of “calm throughout the storm”. This project was not an easy feat but THANK YOU for sticking with me and cheering me on each step of the way. I would also like to thank Dr. Mark Terjesen and Dr. Raymond Giuseppe. You are both brilliant psychologists, and the opportunity to learn from you both has helped me to develop into a thoughtful, caring, and well-rounded clinician. Additionally, thank you Dr. Kate Walton and Nancy Weinstien for sharing your insights and working with me to complete this project.

To my parents, thank you for providing me with the opportunity to follow my dreams. Without your support, guidance, and love, none of this would be possible and I am forever grateful. Johnny, Caroline, and Vicky, thank you all for providing the comic relief necessary to keep daily life interesting, and always picking me up when I am down. Johnny, your contribution to this dissertation, really added the icing on the cake. Uncle Charles, thank you for providing me with a place to complete this project when COVID-19 stopped me from working at Starbucks, you are the best (and “VERY SMART”).

Joe, I cannot possibly thank you for everything you have done for me. Thank you for being my rock and always sticking with me throughout this roller coaster ride. I couldn't have picked a better partner and I cannot wait to see where life takes us.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	ii
LIST OF TABLES.....	vi
INTRODUCTION.....	1
Statement of the Problem.....	1
LITERATURE REVIEW.....	3
College Admissions Testing.....	3
Test Optional Admissions.....	3
Personality and Academic Achievement.....	5
Five-Factor Model and University Achievement.....	6
Self-Report Limitations.....	9
ACT® Tessaera®.....	12
Present Study.....	14
CHAPTER 1.....	15
Hypotheses.....	15
Hypothesis 1a Reliability	15
Hypothesis 1b Internal Structure Validity.....	15
Hypothesis 1c Construct Validity.....	16
Hypothesis 1d Convergent and Discriminant Validity.....	16
Hypothesis 1e Incremental Validity.....	16
Method.....	16
Participants and Procedures.....	16
Measures.....	19

ACT® Tessera® College Pilot.....	19
Analyses.....	21
Results.....	22
Reliability.....	22
Internal Structure Validity.....	23
Convergent and Discriminant Validity	27
Self-Report x Situational Judgement Tests.....	28
Self-Report x Forced Choice.....	28
Situational Judgement Tests x Forced Choice.....	28
Test-Criterion Validity	29
Incremental Validity.....	30
Discussion.....	30
Likert-Items.....	31
Situational Judgement Tests.....	33
Forced-Choice.....	35
Convergent and Discriminant Validity	35
Predictive Validity.....	36
Incremental Validity.....	37
General Conclusions and Considerations.....	38
Limitations and Directions for Future Research	38
CHAPTER 2.....	40
Hypotheses.....	40
Method.....	40

Participants and Procedures	40
Measures: Mindprint Learning®	42
Analyses.....	43
Results.....	44
Regression 1.....	44
Regression 2.....	45
Regression 3.....	45
Regression 4.....	46
DISCUSSION.....	47
Limitations and Future Directions.....	50
IMPLICATIONS FOR SCHOOL PSYCHOLOGISTS.....	52
Improving Applied Practice.....	52
Advocacy.....	52
Graduate Admissions	53
APPENDICES.....	55
Appendix A: Initial Participant Contact	55
Appendix B: Study 1 Formal Invitation to Participate.....	56
Appendix C: Study 1 Informed Consent Agreement.....	57
Appendix D: Study 1 Reminder E-Mail.....	59
Appendix E: Study 2 Invitation to Participate.....	60
Appendix F: Study 2 Informed Consent.....	61
Appendix G: Study 2 Reminder E-Email.....	63
REFERENCES.....	64

LIST OF TABLES

Table 1. <i>Demographic Characteristics (Study 1)</i>	18
Table 2. <i>ACT® Tessera® Definitions</i>	19
Table 3. <i>Self-Report Confirmatory Factor Analysis: Factor Loadings</i>	24
Table 4. <i>Situational Judgment Test Confirmatory Factor Analysis: Factor Loadings</i>	26
Table 5. <i>Multi-Trait Multi Method Matrix Comparing ACT® Tessera® Methods</i>	27
Table 6. <i>ACT® Tessera® Scale Correlations with Grade Point Average</i>	29
Table 7. <i>Incremental Validity ACT® Tessera® Methods:</i>	30
Table 8. <i>Demographic Characteristics (Study 2)</i>	42
Table 9. <i>Hierarchical Regression: Intelligence, Standardized Tests, Social-Emotional Traits (n=38)</i>	45
Table 10. <i>Hierarchical Regression: Standardized Tests, Social-Emotional Traits (n=433)</i>	45
Table 11. <i>Hierarchical Regression: Social-Emotional Traits, Standardized Tests (n=433)</i>	46
Table 12. <i>Hierarchical Regression: College GPA, Standardized Exams, Social Emotional Traits</i>	47

INTRODUCTION

Statement of the Problem

Each year, millions of high school students take standardized college admissions exams, such as the ACT® and SAT® (ACT®, 2020; The College Board, 2020). Students take these exams hoping to demonstrate their ability to perform at the collegiate level and gain admission to their desired schools. Students devote substantial time, money, and energy preparing for these exams, hoping to bolster their scores and increase their likelihood of acceptance (Robinson, 2019). Despite these exams' popularity, there is a growing movement to abandon them in the college admissions process, known as the test-optional movement (FairTest, n.d.). Proponents of this movement highlight standardized testing weaknesses, such as limited exam predictive validity, poor accessibility for disadvantaged populations, and wasted student resources to bolster their argument (Galla et al., 2019; Hoxbey & Turner, 2015; Keiser, et al., 2016). By highlighting these weaknesses, the test-optional movement draws attention to the potential adverse impacts that utilizing standardized exams in college admissions has on students and underscores the importance of re-evaluating current admissions protocols. Therefore, this study sought to investigate the utility of standardized test scores and to identify alternatives to this long-standing admissions practice.

A separate but related line of research has demonstrated interest in how social-emotional characteristics (i.e., noncognitive, personality, social-emotional) contribute to academic success (Sanchez-Ruiz, et al., 2016). Among the studied social-emotional factors, the Five-Factor Model (FFM) of personality is a commonly studied framework, and there is evidence that some FFM traits significantly predict academic success

(Poropat, 2009). In light of this preliminary evidence, FFM traits may offer a potential alternative to using standardized college entrance examinations in college admissions. Therefore, research must investigate the integrity of using social-emotional traits to predict university success, and this dissertation works to propel this line of inquiry forward.

This dissertation's objectives were two-fold, with the ultimate goal of contributing to the field's knowledge of how social-emotional traits contribute to university success. First, study one investigated the psychometric properties of ACT® Tessera®, a recently developed measure designed to assess social-emotional learning in a university population. ACT® Tessera® aligns with the FFM of personality and uses three different response methods to mitigate bias (ACT®, 2018). Given the FFM's evidenced relations with university success, this measure has potential implications for measuring a student's likelihood of university success. Various statistical analyses were used to evaluate the measure's reliability and validity, and the findings were considered alongside previous research. Discussion and conclusions surrounding the scales' strengths and areas for improvement are offered.

Study two evaluated social-emotional traits measured by ACT® Tessera® alongside traditional college achievement predictors, including standardized testing and intelligence. Analyses specifically sought to evaluate if ACT® Tessera®'s social-emotional traits provided incremental validity beyond intelligence and standardized exam scores. The results of this analysis contribute to the discussion on the test-optional movement. A dialogue is presented regarding the importance of advancing this line of study. Implications for school psychologists are discussed.

LITERATURE REVIEW

College Admissions Testing

Historically, college and university admissions committees have relied heavily on standardized testing measures to inform their decisions. Each year, millions of high school students take standardized exams, such as the SAT® and ACT®, and submit their scores from these exams to the colleges and universities they hope to attend (ACT®, 2020; The College Board®, 2020). Although the SAT® intends to measure aptitude and the ACT® intends to measure achievement, they bear striking similarities. Universities consider both exams useful predictors of a student's early college success, and empirical studies report that both exams highly correlate with general intelligence, or *g* (SAT® $r=.72-.86$; ACT® $r=.61-.7$) (Frey & Detterman, 2004; Koenig, Frey, & Detterman, 2008; Syverson, 2007). While colleges evaluate additional student data in conjunction with standardized test performance (i.e., high school grade point average [GPA], letters of recommendation, and writing samples), these exams provide the only standardized data available. Therefore, committees regard these measures as student success predictors with reduced influence of different educational experiences (i.e., course difficulty, course grade inflation) and expect them to facilitate fair comparisons between students across the country (Syverson, 2007). However, despite the long-standing use of these exams, a growing movement encourages colleges and universities to abandon standardized college admissions testing.

Test-Optional Admissions

The National Center for Fair and Open Testing (FairTest) frequently cites problems with using the ACT® and SAT® in college admissions to encourage

universities to reconsider using these exams (FairTest, n.d.). While the movement began at a few small liberal arts colleges in the 1960s, many more colleges have adopted test-optional policies since this time (Lucido, 2018). Before the COVID-19 pandemic, 1,070 colleges and universities were test-optional or test-blind, and this number has increased to over 1,685 for the Fall 2021 semester (FairTest, 2020). Therefore, the test-optional movement has certainly gained traction, and it is essential to consider how adopting these practices may influence the admissions process.

To support the adoption of test-optional policies, proponents identify various problems with using standardized testing in admissions. One argument focuses on efficiency. Over time, various studies have indicated that the SAT® and ACT® are successful predictors of subsequent university performance (Noble & Sawyer, 2002; Sackett, et al., 2009). However, recent evidence finds that these exams' predictive power beyond high school GPA may be meager (Galla et al., 2019). Given that these tests are expensive, high stress, and time-consuming, if admissions committees can accurately predict college performance without them, it may be helpful to consider alternatives to this common practice for the sake of allowing students to allocate their time and energy to more meaningful activities.

In addition to efficiency, student-imposed barriers are essential to consider. For example, students may limit their applications to schools they meet or exceed the university reported average standardized test scores. Unfortunately, qualified candidates may be reluctant to apply to a school if they do not meet their average standardized test scores, even if they have other desirable qualifications (Lucido, 2018). These self-limiting decisions may be especially harmful, given research that finds that students from

low-income families frequently underestimate their ability (Lucido, 2018; Hoxbey & Turner, 2015). Therefore, test-optional admissions may prevent candidates from applying to schools they are qualified to attend and may limit the diversity of students who attend these schools. On the contrary, test-optional policies encourage students from diverse cultural backgrounds to apply to universities that they may have previously avoided if required to submit standardized exam scores (Lucido, 2018).

In addition to self-inflicted barriers to college admissions, standardized tests may limit a college admissions committee's decision-making process. Standardized testing may encourage hurried and limited decision-making. A university may ignore a student's application if they fail to reach a particular profile put forth by the school or make quick judgments about an applicant based on their test score (Lucido, 2018). Given the evidence to suggest that standardized test scores under-predict the achievement of particular populations (i.e., women, black students) (Keiser, et al., 2016; Lawlor, Richman, & Richman, 1997; Shewach, et al., 2017), these practices may lead to a biased admissions protocol. Given these limitations, numerous universities have abandoned using standardized tests in their admissions process, and researchers have begun looking for alternative ways beyond intelligence estimates to improve academic success predictions (Syverson, 2007).

Personality and Academic Achievement

By recognizing standardized testing's limits to selecting quality candidates, several researchers have investigated social-emotional factors to understand their academic success relationship. Notably, researchers are often interested in studying personality as a social-emotional influence on achievement (Sanchez-Ruiz, et al., 2016).

Studies have established personality as a predictor of academic achievement, and this relationship remains significant when controlling for intelligence (Bratko, Chamorro-Premuzic, & Sacks, 2006; Nofle & Robins, 2007; Ziegler, et al., 2010). Additionally, some studies even suggest that personality may predict academic success more strongly than intellectual ability (Conard, 2006; Furnham & Chamorro-Premuzic, 2004; Furnham, Chamorro-Premuzic, & Mcdougall, 2003). Given these findings, research should continue exploring the relationship between personality and university achievement to determine if they can improve academic success predictions.

Five-Factor Model and University Achievement.

The Five-Factor Model (FMM) is a widely accepted personality framework that researchers use to conceptualize personality (Laidra, Pullmann, & Allik, 2007; Novikaova & Vorobyeva, 2017; Poropat, 2009). The FFM taxonomy emerged from factor analyses of English words describing personality traits, which repeatedly yielded five general personality factors that subsumed all of the studied traits (John & Srivastava, 1999). These five factors are labeled: Conscientiousness, Openness to Experience (also referred to as Openness), Agreeableness, Extraversion, and Neuroticism (Digman, 1990). Considering the FFM's empirical backing, the present study used this model to conceptualize social-emotional learning. The following discussion briefly reviews FFM traits and their association with university achievement.

Studies investigating university achievement and the FFM have reported varying predictive validity across traits (Poropat, 2009). Conscientious individuals are disciplined, dutiful, and achievement-oriented; whereas, individuals low on this trait may be considered irresponsible, negligent, or carefree (Trull & Widiger, 2013). Regarding

Conscientiousness' relationship with academic achievement, O'Connor and Paunonen (2007) posit that "Conscientiousness is clearly an important determinant of academic success..." (pp. 976), and a myriad of studies support this claim (Poropat, 2009).

Researchers have established a positive relationship between Conscientiousness and university achievement across various outcome variables (i.e., exam performance, Grade Point Average (GPA), overall course performance; Busato, et al., 1999; Conard, 2006; Paunonen & Ashton, 2001) and across time (i.e., freshman to senior year; Wagerman, & Funder, 2007). Evidence also suggests that Conscientiousness is just as valuable as intelligence when comparing the correlations between Conscientiousness and intelligence with university achievement (Poropat, 2009). Therefore, the overwhelming evidence establishing Conscientiousness' importance to university achievement suggests that this trait may be beneficial to consider when predicting achievement.

While research has demonstrated a positive relationship between Conscientiousness and university achievement, claims about the other FFM traits such as Openness and Extraversion are more ambivalent. Individuals who display high Openness levels are curious, imaginative, artistic, unconventional, and sensitive to art and beauty (Costa & McCrae, 1992; John & Srivastava, 1999). There is empirical support linking Openness to numerous learning variables important for academic success, such as intelligence (McCrae, 1987) and the desire to seek out learning opportunities (Fiske 1949). However, investigations that specifically investigate Openness in university performance report negligible associations between achievement measures and Openness (Chamorro-Premuzic & Funham, 2003; Duff, et al., 2003; Gatzka & Hell, 2018; O'Connor & Paunonen, 2007; Trapmann, et al., 2007). To resolve the empirical discord,

researchers propose that confounding variables (Farsides & Woodfield, 2003; O'Connor & Paunonen, 2007) or subordinate openness factors may moderate the relationship between academic achievement and Openness (Gatzka & Hell, 2018). As such, the relationship between Openness and university achievement is unclear and more research is required to understand the relationship between these variables.

Like the relationship between Openness and academic achievement, the relationship between Extraversion and academic achievement is equivocal. Extraversion describes individuals who are sociable, active, desire interpersonal experiences, and frequently experience positive emotions (Costa & McCrae 1992; Wilt & Revelle, 2017). When considering the relationship between Extraversion and academic achievement, some studies report a slight negative correlation between these variables (Finlayson, 1970; Kline, 1966; Nechita, et al., 2015; O'Connor & Paunonen, 2007). However, several meta-analytic studies fail to detect a meaningful relationship between these variables (Poropat, 2009; Trapmann, et al., 2007). Therefore, although there is some evidence to suggest that Extraversion is related to academic achievement in a university setting, more research is required before making any strong claims about its relation to academic success.

While there is evidence that some FFM factors predict university achievement, other factors are unrelated to achievement. Specifically, Neuroticism, or an individual's tendency to experience psychological distress (Costa & McCrae, 1992), yields little influence on performance (O'Connor & Paunonen, 2007; Poropat, 2009; Trapmann, et al., 2007). Similarly, when considering Agreeableness, a factor reflective of an individual's interpersonal behavior (i.e., agreeable individuals are trusting, sympathetic,

and cooperative) (Costa & McCrae, 1992), empirical evidence suggests that this variable is unimportant for undergraduate achievement (O'Connor & Paunonen, 2007; Poropat, 2009; Trappmann, et al., 2007).

Although the predictive validity of FFM traits and university achievement is variable, evidence suggests that personality is a significant social-emotional university achievement predictor (Poropat, 2009). Therefore, it is crucial to advance our understanding of these relationships through continued research efforts. Given that the college admissions process generally ignores social-emotional constructs when predicting candidates' university achievement potential, it is critical to investigate these constructs further to determine if considering these variables in university admissions decisions would improve predictions.

Self-Report Limitations

Given the evidence that social-emotional traits, such as the FFM facets, may improve academic achievement predictions, researchers must evaluate ways to measure these constructs. The findings reported above regarding FFM and achievement generally rely on traditional, self-reported personality assessment methods to measure traits (i.e., NEO-PI-R, The Big Five Inventory, the Sixteen Personality Factor Questionnaire; Cattell & Mead, 2008; Costa & McCrae, 1992; John & Srivastava, 1999; Trappmann, et al., 2007). Although traditional self-report measures are efficient (i.e., effective, fast, and inexpensive) and provide insight regarding an individual's self-perception, they are vulnerable to method bias (Paulhus & Vazire, 2007; Vazire & Mehl, 2008). Therefore, before the college admissions process begins considering personality measures, it is

essential to understand method bias, consider how it affects personality assessment, and reduce such effects in the measurement methods used.

Method bias is any score variation reported by a measure attributable to factors external to actual differences in the intended variable or random error (Campbell & Fiske, 1959). The most widely studied form of method bias concerning personality assessment is called “response sets.” Response sets refer to an individual’s tendency to systematically complete items, which reduces the scale’s validity. Three common response sets are acquiescent responding, socially desirable responding, and extreme responding. Socially desirable responding (SDR) describes patterns produced when individuals inaccurately portray themselves because they are concerned with how others will receive their responses (e.g., exaggerating, faking, lying). Acquiescent responding occurs when respondents report high or low proportions of the same response, regardless of item content (i.e., strongly agreeing with all items). Extreme responding refers to response styles in which respondents repeatedly select the extreme responses on a scale (i.e., only reporting 1’s or 7’s on a 7-point Likert scale). Although there are many method bias sources, the preceding discussion illustrates some of the ways that self-report personality measures are vulnerable to bias (Paulhus & Vazire, 2007).

Although some method bias studies suggest that response sets yield negligible influence on assessment results, substantial, conflicting literature suggests that such biases meaningfully impair scale interpretation and can reduce the quality of decisions made based on data generated from these measures (Christiansen, et al., 1994; Mueller-Hanson, Heggstad, & Thornton, 2003). Generally, Jackson and Messick (1958) cite evidence that response sets (i.e., stylistic determinants) account for a large proportion of

variance on personality scales; thus, reducing such scales' content validity and interpretability. Further, research investigating particular types of method bias (i.e., acquiescent responding, SDR) finds significant biases in factorial structure or distortions in the intended measure, which diminish scale validity (e.g., content and criterion) (Danner, Aichholzer, & Rammsedt, 2015; McFarland & Ryan, 2000; Mueller-Hanson, Heggstad & Thorton, 2003; Rammstedt & Farmer, 2013; Rammsedt, Kemper, & Borg, 2013). Specifically, Niessen, Meijer, & Tenderio (2017) reported that when participants were applying for admission to a university, social-emotional measures, including Conscientiousness, were inflated compared to a low-stakes scenario (i.e., research setting), and these effects attenuated predictive and incremental validity. The studies mentioned above caution personality assessment consumers to consider method bias's effect on the accuracy of assessment before using the data to make critical decisions.

Researchers and clinicians have long recognized method bias's potential for adversely influencing assessment. Thus, they have employed many techniques to ameliorate these effects. For example, some assessment tools possess validity scales to identify socially desirable responses (e.g., NEO-PI-R Positive Presentation Management scale) (Schinka, Kinder, & Kramer, 1997), and other researchers use statistical methods to control for the effects of acquiescent responding (McCrae, Herbst, & Costa, 2001; Rammstedt, & Farmer, 2013). However, these method-bias reduction methods are reactive as opposed to proactive. Individuals complete the scale *first*; then, if these methods detect bias, the data is deemed invalid or adjusted. Such methods are inefficient, wasting both the researcher and the individual's resources (i.e., time, materials).

Alternatively, McCrae (2018) advocates combating method bias by requesting multiple

informants' complete personality inventories. However, in practice (i.e., school, occupational settings), multiple informants may be inaccessible. Therefore, developing alternative ways to combat method bias is necessary to improve current personality assessment methods.

Recognizing the need for improved personality assessment, several researchers have made recommendations to facilitate such developments. Notably, Funder (2002) suggested that “personality psychology’s methods will need to expand into innovative techniques that go beyond, without replacing, self-report measures...” and that personality assessment should consider situational and behavioral variables more regularly to improve measurement. Furthermore, McDonald (2008) proposed that multiple personality assessment methods be combined when measuring personality. The author supports her position, stating that using multiple assessment methods will lead to richer, more valid, and more informative measurements. Together, these suggestions imply that a multi-method assessment tool that combines self-reports with situational and behavioral measures might improve personality measurement. Although these recommendations are over a decade old, a personality assessment reflecting these recommendations was unavailable until recently.

ACT® Tessera®

ACT® Tessera® Social and Emotional Learning Assessment System offers promising improvements to personality assessment. ACT® Tessera® strives to measure social and emotional skills through the FFM framework's lens to guide intervention programs and promote school, career, and life success (ACT®, 2018). Notably, the FFM Conscientiousness, Extraversion, Openness, Neuroticism, and Agreeableness are

reflected in the ACT® Tessera® Grit, Leadership, Curiosity, Resilience, and Teamwork scales, respectively. ACT® first published a school-aged version of ACT® Tessera® in 2018, with data supporting its use for children in grades 6-12. Recently, ACT® has adapted this social-emotional assessment system to measure these constructs in university students. However, empirical studies have not yet established the measures' reliability and validity. Contrary to existing personality assessment measures, ACT® Tessera® uses *multiple response methods*, including 40 self-report, Likert scale items (SR Likert), 30 forced-choice (FC) items, and 30 situational judgment test (SJT) items to measure social-emotional learning skills.

ACT® Tessera®'s utilization of FC and SJTs in conjunction with traditional self-report methods strives to improve traditional, mono-method personality measures. Specifically, both SJTs and forced-choice formats are considered more resistant to "faking good" than self-report Likert-style measures (Olaru, et al., 2019). Remarkably, when comparing FC methods with other assessment methods, studies suggest that forced-choice methods may be more resistant to socially desirable responses when respondents are highly motivated than more commonly used single stimulus traditional self-report personality items (Christiansen, Burns, & Montgomery, 2005). Additionally, Bartram (2007) reported that studies that use a forced-choice methodology resulted in a 50% increase in criterion validity compared to a normative instrument.

Similarly, several studies support SJT's ability to validly assess personality (Mussel, Gatzka & Hewig, 2016). Specifically, Lievens & Coetsier, (2003) found that when using SJTs to complement other admissions information, they were better predictors of achievement (i.e., first-year course average) than other predictors (i.e.,

intellectual ability). Additionally, Cousans et al., (2017) provided solid predictive validity for using SJTs in medical school admissions when combining SJTs with multiple assessment measures. Therefore, considering the validity of SJTs and FC methods for assessing personality and predicting academic performance, ACT® Tessera® likely provides a valid measure of personality that can be used to make an informed educational decision.

Present Study

Colleges and universities are increasingly abandoning standardized testing in the college admissions process, creating a need to investigate alternatives to this practice. Given the evidence to suggest that social-emotional skills can successfully predict university outcomes, an exciting and necessary area for investigation emerges. The development of ACT® Tessera® offers promising improvement to social-emotional skill measurement and, potentially, the college admissions process. This two-study dissertation will first investigate the psychometric properties of ACT® Tessera's® college version to evaluate its ability to measure FFM traits validly and reliably. Then social-emotional skills' abilities to make university success predictions will be considered alongside traditionally used predicative variables. Together, these studies intend to make significant contributions to the field by validating a scale in line with a popular, empirically supported framework and considering if the scale can aid in improved educational decision making. Universities, educators, psychologists, and test developers can all benefit from this work.

CHAPTER 1

Study one explored ACT® Tessera®'s ability to validly and reliably measure social-emotional skills through the FFM framework in an undergraduate population. This study specifically evaluated the scale's reliability, internal structure validity, construct validity, convergent validity, and incremental validity.

Hypotheses

Given the evidence provided in empirical studies, which demonstrated that SR Likert, SJT, and FC assessment methods could accurately measure social-emotional traits, I expected ACT® Tessera® to provide valid and reliable FFM measures (Goldberg, 1992; Costa & McCrae, 2008; O'Neill et al., 2017; Olaru et al., 2019). More specifically, this study investigated the following hypotheses:

Hypothesis 1a Reliability

I predicted that SR and SJT scales would yield "acceptable" ($\alpha = .06-.07$) or "good" ($\alpha = .08$ or higher) internal consistencies (George & Mallery, 2003). However, I expected FC reliability analyses to yield lower Cronbach's alpha levels than the SR and SJT scales while still demonstrating "acceptable" internal consistencies (Saville & Willson, 1991).

Hypothesis 1b Internal Structure Validity

ACT® Tessera® was developed to align with the FFM (ACT®, 2018), and support for this model is well documented (Laidra, Pullmann, & Allik, 2007; Novikaova & Vorobyeva, 2017; Poropat, 2009). Therefore, I expected the ACT® Tessera® SR and SJT items to fit a five-factor CFA specifying each social-emotional trait as a factor.

Hypothesis 1c Construct Validity

Regarding construct validity, I predicted that ACT® Tessera® Grit (FFM Conscientiousness) measures would demonstrate significant positive correlations with academic achievement. Contrarily, I expected ACT® Resilience and Teamwork (FFM Neuroticism and Agreeableness, respectively) to yield negligible correlations with achievement (Poropat, 2009).

Hypothesis 1d Convergent and Discriminant Validity

I predicted that the same factor measured by different methods would highly correlate. However, I expected that the different factors and different methods would yield low correlations.

Hypothesis 1e Incremental Validity

Given evidence reported in the literature review, which illustrated SR methods vulnerability to bias (Danner, Aichholzer, & Rammsedt, 2015; McFarland & Ryan, 2000; Mueller-Hanson, Heggstad & Thorton, 2003), I expected the FC and SJT scales to improve academic success predictions above and beyond SR predictions.

Method

Participants and Procedures

Participants were recruited through email communications. All undergraduate students at St. John's University (Queens, NY campus) were initially contacted with an email alerting them to the study (Appendix A; $N=10,255$). The following day, a second email was sent to all students, formally inviting them to participate in the study (Appendix B). The second email provided a detailed study description and a hyperlink that directed students to the Qualtrics survey platform. Students were electronically

presented an IRB-approved consent form (Appendix C), then directed to the ACT® Tessera® assessment items (ACT® Tessera®, 2017). As an incentive to participate in the study, all participants were allowed to request their survey results and were automatically entered into a raffle to win one of two \$50 Amazon gift cards. To maximize participation, email reminders were sent to all potential participants who did not complete the survey 5 and 7 days after sending the initial email (Appendix D). The survey remained accessible to participants for two weeks. When the survey closed, the Office of Institutional Research at St. John's University provided the following data for each participant: student class year (e.g., Freshman, Sophomore, Junior, Senior), age, sex, ethnicity, high school average, SAT®-Verbal score, SAT®-Math score, ACT® score, college GPA, College Major, and honors student status.

562 St. John's University (Queens, New York Campus) undergraduate students comprised the final sample ($n=562$). 706 people opened the survey. However, 44 respondents were excluded from analyses for failing to provide valid consent. Another 44 responses were removed because the respondent initiated multiple survey attempts. In these cases, a respondent's first complete response was retained for analysis. 27 cases were excluded for providing incorrect responses to attention check items. 2 cases were deleted due to low variability ($<.01$), and 1 case was deleted for extreme variability (>4.0). 26 cases were deleted for failing to respond to all items. Participants ranged from age 18 to 35, ($M=19.97$; $SD=1.88$). 110 freshmen (19.6%), 137 sophomores (24.4%), 134 juniors (23.8%), and 181 seniors (32.2%) participated in study one. 21.7% percent of participants were honor students. The sample was 67.8% female. Participants identified as White (40.6%), Asian (17.8%), Black/African American (17.4%), Hispanic (16.5%),

American Indian or Alaska Native (.4%), Native Hawaiian/Pacific Islander (.2%). 4.8% identified with two or more races, and 2.3% indicated an unknown ethnicity.

Demographic characteristics of study 1 participants appear in Table 1.

Table 1

Demographic Characteristics (Study 1)

	N	Percentage (%)
Gender		
Female	381	67.8
Male	181	32.2
Age		
18	118	21
19	123	21.9
20	130	23.1
21	125	22.2
22	44	7.8
23	7	1.2
24	7	1.2
25 and older	8	1.4
Class Standing		
Freshman (1 st year)	110	19.6
Sophomore (2 nd year)	137	24.4
Junior (3 rd year)	134	23.8
Senior (4 th year)	181	32.2
Ethnicity		
Two or more races	27	4.8
American Indian or Alaska Native	2	.4
Asian	100	17.8
Black or African American	98	17.4
Hispanic	93	16.5
Native Hawaiian or Other Pacific Islander	1	.2
White	228	40.6
Unknown	13	2.3
Honors Status		
Honors Student	112	21.7
Non-Honors Student	440	78.3

Note. N=562

Measures

ACT® Tessera® College Pilot

ACT® Tessera® is a multi-trait multi-measure assessment system that assesses five social-emotional learning skills that reflect FFM factors (ACT®, 2018). Specifically, *Grit*, *Teamwork*, *Resilience*, *Curiosity*, and *Leadership* can be understood as FFM Conscientiousness, Agreeableness, Emotional Stability, Openness, and Extraversion, respectively. For social-emotional trait definitions, refer to Table 2.

Table 2

ACT® Tessera® Definitions

ACT® Tessera® Social-Emotional Skill	Definition
Grit	The extent to which a student's actions demonstrate goal striving, dependability, and attention to detail at school.
Teamwork	The extent to which a student's actions demonstrate collaboration, empathy, helpfulness, trust, and trustworthiness.
Resilience	The extent to which a student's actions demonstrate stress management, emotional regulation, a positive response to setbacks, and poise.
Curiosity	The extent to which a student's actions demonstrate creativity, inquisitiveness, flexibility, open-mindedness, and embracing diversity.
Leadership	The extent to which a student's actions demonstrate assertiveness, influence, optimism, and enthusiasm.

ACT® Tessera® uses three methods to assess social-emotional skills: self-report Likert (SR Likert) items, forced-choice (FC), and situational judgment tests (SJTs). First, the 40 self-report Likert items request that individuals read items and indicate the degree to which they agree with the statements on a 6-point Likert scale (strongly agree, agree,

somewhat agree, somewhat disagree, disagree, strongly disagree). A SR Likert example is as follows: “I finish homework assignments before they are due.” Two negatively phrased items per scale were reverse-scored, and then the scale score for each social-emotional learning skill was derived by taking the mean score of the six items per scale.

Thirty forced-choice items, arranged into ten triads, were administered. The ten triads accounted for every possible combination of three traits. Participants were presented three statements and then selected the statement they identified most strongly with by selecting “most like me” and the statement they identified least with by selecting “least like me.” One statement in each triad was not selected. A sample forced-choice triad is as follows: “I do more than what my teachers expect,” “I am concerned about other students,” and “I cope well with stressful assignments.” Ipsative scores were calculated for all forced-choice items by creating rank-ordered scores for each triad (Most like me=3, Not selected=2, least like me=1). Scale scores were then calculated by calculating the mean score from how the participant ranked the six individual items per scale. Again, negatively phrased items were reverse-scored and incorporated into the mean.

Lastly, Situational Judgment Tests (SJTs) presented participants with hypothetical situations and potential behavioral responses to each situation. Participants indicated the likelihood that they would demonstrate each behavioral response using a six-point Likert scale (very unlikely, unlikely, may or may not, likely, very likely). Participants were presented with ten situations (two per skill) and three behavioral responses for a total of thirty SJT items. Each item independently contributed to the individual’s SJT score. To score, SJT’s responses were first correlated with the individuals’ self-reported Likert

responses. Items that correlated negatively with self-report items were reverse-scored before calculating the SJT score. Scale scores were derived by generating mean scores items per skill.

Analyses

To evaluate this study's hypotheses, I conducted a series of statistical analyses. First, to evaluate *reliability*, I calculated Cronbach's alpha (α) for each SR Likert, FC, and SJU scale to measure internal consistency. All Cronbach's alpha calculations were conducted in SPSS Version 26 (IBM, 2019). Then, to evaluate *internal structure validity*, I ran two Confirmatory Factor Analyses (CFA) to test the SR Likert and SJT model fit in MPlus8 (Muthén & Muthén, 2007). Each CFA specified the five ACT® Tessera® social-emotional skills as factors and used weighted least squares estimation. Model fit was assessed with model fit indices, including the Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Residual (SRMR). Additionally, factor loadings were examined to evaluate further if the items accurately loaded on the intended trait.

To assess ACT® Tessera®'s *convergent and discriminant validity*, I correlated the 15 scale scores and created a multi-trait multi-method correlation matrix. Correlations between scales that utilized different methods to measure the same social-emotional traits were examined and expected to demonstrate higher correlations than measures that intend to measure different traits.

To assess *test-criterion validity*, I calculated correlations between college GPA and the ACT® Tessera® scale scores and aggregate scores and compared them to the empirical findings.

Finally, to assess *incremental* validity, I conducted a hierarchical regression and evaluated the change in college GPA prediction given the introduction of additional measurement methods beyond self-report.

Results

Reliability

To assess ACT® Tessera®’s scale reliability, Cronbach’s alpha for each of the SR Likert, SJT, and FC scales was calculated. Internal consistency for each FFM SR Likert scale was considered either “acceptable” or “good” (George & Mallery, 2003). The alpha values for the SR Report Likert scales were: Leadership $\alpha = 0.81$, Teamwork $\alpha = 0.76$, Grit $\alpha = 0.86$, Resilience $\alpha = 0.72$, and Curiosity $\alpha = 0.77$.

Cronbach’s alpha for each FFM SJT scale was also calculated. Cronbach’s alpha values for the Leadership and Grit scales were “acceptable,” and the alpha values for the Teamwork, Resilience, and Curiosity scales were “questionable” (George & Mallery, 2003). Cronbach’s alpha for each FC scale was as follows: Leadership $\alpha = 0.73$, Teamwork $\alpha = 0.65$, Grit $\alpha = 0.72$, Resilience $\alpha = 0.69$, and Curiosity $\alpha = 0.62$.

Lastly, Cronbach’s alpha for each FC scale was also calculated. Internal consistency for the Leadership, Grit, Resilience, and Curiosity scales was “poor,” and the internal consistency for the FC Teamwork Scale was “acceptable” (George & Mallery, 2003). The Cronbach’s alpha values were as follows: Leadership $\alpha = 0.56$, Teamwork $\alpha = 0.36$, Grit $\alpha = 0.54$, Resilience $\alpha = 0.56$, and Curiosity $\alpha = 0.54$. Notably, these findings should be interpreted cautiously, given known limitations regarding assessing ipsative scale reliability. For more information regarding these limitations, refer to the discussion section below.

Internal Structure Validity

To assess internal structure validity, I ran a CFA evaluating the SR Likert data's fit to the FFM. The model specified five factors, Grit, Teamwork, Reliance, Curiosity, and Leadership. The analysis used weighted least squares estimation because this estimation method was created explicitly for ordinal data (i.e., Likert data). The Root Mean Square Error of Approximation (RMSEA) for the SR CFA model was .07 [CI=.06-072]. While low RMSEA values are most desirable, Browne and Cudeck (1993) consider values within the .05 to .08 range to represent a "fair" fit. Similarly, the SRMR index was .07, which is below .08, and suggests a good fitting model (Hu & Bentler, 1999; SRMR = .07). However, Hu and Bentler's (1999) model fit criteria, suggests that the CFI and TLI should be close to .95 to make confident assertions regarding satisfactory model fit. Therefore, the CFI and TLI, in this case, suggests a "poor" fitting model [CFI= .84; TLI=.83]. Standardized and unstandardized factor loadings are presented in Table 3.

Table 3*Self-Report Confirmatory Factor Analyses: Factor Loadings*

	B	β	SE	z
Teamwork 1	1.00	.63	--	--
Teamwork 2	.95	.60	.08	12.13
Teamwork 3	.93	.59	.08	12.49
Teamwork 4	.98	.62	.08	11.88
Teamwork 5	1.16	.73	.08	14.58
Teamwork 6	1.07	.68	.08	13.69
Teamwork 7	1.11	.70	.08	14.83
Teamwork 8	1.01	.64	.08	13.40
Leadership 1	1.00	.75	--	--
Leadership 2	.92	.69	.05	17.88
Leadership 3	.90	.68	.05	16.92
Leadership 4	.66	.50	.06	11.79
Leadership 5	.62	.46	.05	11.54
Leadership 6	.87	.65	.05	18.42
Leadership 7	1.10	.74	.05	16.79
Leadership 8	1.02	.60		
Resilience 1	1.00	.56	--	--
Resilience 2	1.15	.64	.10	11.78
Resilience 3	.89	.50	.08	11.00
Resilience 4	.88	.49	.09	9.50
Resilience 5	.99	.56	.09	11.21
Resilience 6	.86	.48	.09	9.17
Resilience 7	1.14	.64	.11	10.40
Resilience 8	.93	.52	.10	9.50
Curiosity 1	1.00	.56	--	--
Curiosity 2	.96	.54	.09	10.53
Curiosity 3	1.06	.59	.09	12.23
Curiosity 4	.99	.56	.09	11.60
Curiosity 5	1.19	.67	.08	14.92
Curiosity 6	1.22	.69	.10	12.34
Curiosity 7	1.25	.70	.10	12.84
Curiosity 8	1.20	.67	.10	12.60
Grit 1	1.00	.82	--	--
Grit 2	.85	.70	.04	23.89
Grit 3	.92	.76	.03	27.50
Grit 4	.71	.58	.04	16.24
Grit 5	.87	.72	.04	23.53
Grit 6	.79	.66	.05	17.70
Grit 7	.85	.70	.04	22.19
Grit 8	1.03	.84	.03	29.98

Note. Standardized (B) and unstandardized(β) factor loadings for SR items with FFM traits specified.

I conducted a second CFA to assess the SJT data's fit to the FFM model. Again, the model specified the five Tessera® social-emotional traits and used weighted least squares estimation. Similar to the SR model, RSMEA and SRMR for the SJT CFA supported the data's fit to the five-factor FFM model (RMSEA=.08 [CI=.07-072]; SRMR = .07; Brown & Cudeck, 1996; Hu & Bentler, 1999). However, the CFI and TLI again suggested questionably fitting data (CFI = .81, TLI =.79). Standardized and unstandardized factor loadings are presented in Table 4.

Table 4*Situational Judgement Test Confirmatory Factor Analyses: Factor Loadings*

	<i>B</i>	<i>B</i>	SE	<i>z</i>
Teamwork 1	1.00	.73	--	--
Teamwork 2	.89	.65	.06	13.96
Teamwork 3	1.07	.78	.07	15.19
Teamwork 4	.58	.43	.07	8.37
Teamwork 5	.58	.42	.07	8.60
Teamwork 6	.30	.22	.07	4.21
Leadership 1	1.00	.46	--	--
Leadership 2	1.45	.67	.14	10.52
Leadership 3	1.53	.71	.13	11.50
Leadership 4	1.40	.64	.14	10.07
Leadership 5	1.60	.73	.14	11.29
Leadership 6	1.58	.73	.14	11.15
Resilience 1	1.00	.58	--	--
Resilience 2	.92	.54	.09	10.07
Resilience 3	.66	.39	.09	7.72
Resilience 4	1.37	.80	.12	11.72
Resilience 5	1.27	.74	.11	11.67
Resilience 6	.60	.35	.09	6.98
Curiosity 1	1.00	.56	--	--
Curiosity 2	1.21	.68	.12	10.30
Curiosity 3	.73	.41	.09	7.97
Curiosity 4	.57	.32	.10	5.60
Curiosity 5	1.22	.68	.13	9.64
Curiosity 6	.682	.38	.10	6.73
Grit 1	1.00	.39	--	--
Grit 2	1.75	.68	.20	8.55
Grit 3	1.39	.54	.17	8.30
Grit 4	1.54	.60	.20	7.53
Grit 5	1.94	.76	.23	8.63
Grit 6	2.10	.82	.25	8.51

Note. Standardized (*B*) and unstandardized (β) factor loadings for SJT items with FFM traits

specified.

Convergent and Discriminant Validity

Correlations among the 15 scale scores were computed and used to generate a Multi-Trait Multi-Method Matrix. The average monotrait-hetromethod correlation was .50, providing evidence for moderate convergent validity.

Table 5

Multi-Trait Multi Method Matrix Comparing ACT® Tessera® Methods

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Likert														
1. Grit	--													
2. Teamwork	.43	--												
3. Resilience	.32	.36	--											
4. Curiosity	.35	.47	.36	--										
5. Leadership	.23	.30	.26	.51	--									
Situational Judgement Tests														
6. Grit	.68	.40	.26	.31	.30	--								
7. Teamwork	.34	.52	.34	.34	.18	.31	--							
8. Resilience	.10	.08	.42	.21	.33	.14	.18	--						
9. Curiosity	.31	.28	.36	.47	.27	.30	.26	.12	--					
10. Leadership	.46	.42	.35	.44	.58	.47	.32	.28	.35	--				
Forced Choice Items														
11. Grit	.67	.32	.39	.23	.24	.50	.21	.23	.20	.46	--			
12. Teamwork	.13	.38	.33	.20	.32	.17	.29	.26	.21	.29	.20	--		
13. Resilience	.24	.14	.55	.12	.22	.19	.06	.45	.15	.30	.52	.38	--	
14. Curiosity	.11	.21	.16	.60	.49	.16	.18	.18	.24	.27	.17	.25	.12	--
15. Leadership	.19	.20	.11	.43	.70	.21	.12	.17	.14	.42	.19	.24	.04	.67

Note. Bolded correlations indicate scales that are intended to measure the same skill and are expected to be greatest.

Self-Report x Situational Judgment Tests

When analyzing the reported correlations between SR and SJT measures of the same trait (i.e., SR Likert Grit and SJT Grit), the average correlation was .53 (ranged from .42-.68). These correlations were generally more robust than those reported between the SR and SJT correlations measures for different factors (i.e., SR Likert Grit and SJT Resilience), which averaged to .30 ($r=.08-.46$). The average mono-trait correlation was significantly higher than the average hetro-trait correlations ($z=4.52$, $p<.01$, two-tailed). However, the SJT Extraversion scales correlated unexpectedly high with the other traits' SR measures ($r=.35-.46$) (Conscientiousness, Agreeableness, Emotional Stability, and Openness).

Self-Report x Forced Choice

The convergent validity between the FC and SR Likert measures was moderate, with the correlations between measures of the same factors averaging to .58 ($r=.38-.70$). There was also good evidence for discriminant validity between FC items and SR Likert items, as the correlations between measures of different factors were generally low, averaging to .24 ($r=.11-.49$). The average mono-trait correlation was significantly higher than the average hetro-trait correlation ($z=6.76$, $p<.01$, two-tailed). However, the correlation between FC Curiosity and SR Likert Leadership scales was unexpectedly high ($r=.49$), as was the relationship between FC Leadership and SR Curiosity ($r=.43$).

Situational Judgment Tests x Forced Choice

Considering the relationship between FC and SJT measures, evidence for convergent validity was variable, and the mono-trait correlations averaged to .38 ($r=.20-.50$). Convergent validity was strongest between FC and SJT Grit ($r=.50$), Resilience

($r=.45$), and Leadership ($r=.42$). However, the convergent validity between FC and SJT Teamwork ($r=.29$) and Curiosity scales ($r=.24$) was weak. On the other hand, evidence for discriminant validity was good with hetero-trait, hetero-method correlations averaging to .22 ($r=.06-.46$). Despite the weaker correlations between FC and SJT measures, the average mono-trait correlation was significantly higher than the average hetro-trait correlation ($z=2.89, p<.01$, two-tailed).

Test-Criterion Validity

Correlations between the 15 scale scores and college Grade Point Average, as reported by the SJU Office of Institutional Research, are reported in Table 6.

Table 6

ACT® Tessera® Scale Correlations with Grade Point Average

	SR	SJT	FC	Aggregate
Leadership	.03	.13*	.05	.09**
Teamwork	.16**	.02	.02	.08**
Grit	.28*	.25*	.26*	.30*
Resilience	-.01	-.05	.04	-.01
Curiosity	.02	.01	.01	.02

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

Consistent with findings reported in the empirical literature and Hypothesis 1c, Grit consistently, significantly, and positively correlated with GPA across all measurement methods and the aggregate score ($r=.30$). Curiosity and Leadership exhibited negligible correlations with academic achievement ($r=.01, r=.02$, respectively). Generally, the Leadership and Teamwork scale also yielded small correlations with achievement. Unexpectedly, the Leadership SJT and Teamwork SR Likert scores significantly correlated with college GPA.

Incremental Validity

A three-stage hierarchical multiple regression was conducted with college GPA as the dependent variable. SR Likert scales were entered into block one. Situational Judgment Test scores were entered into block two. Forced Choice scores were entered at stage three. Table 7 shows model summary statistics and the change in R².

Table 7

Incremental Validity: ACT® Tessera® Methods

	R	R²	Adjusted R²	SE	R² Change
SR Likert	.31	.10	.09	.56	--
SR Likert, SJT	.34	.12	.10	.56	.02*
SR Likert, SJT, FC	.36	.13	.10	.56	.01

Note. *p<.01

Discussion

Study one evaluated ACT® Tessera®'s ability to validly and reliably measure social-emotional skills in a university population. ACT® Tessera® is unique relative to other FFM-based measures available, as it goes beyond using SR Likert scales by including multiple item formats to evaluate social-emotional skills. The scale's multi-method format is attractive, given issues regarding response bias observed in SR Likert measurement (Johnson & Rothstein, 1994; Mueller-Hanson, Heggestad, & Thornton, 2003). This study specifically investigated the SR Likert, SJT, and FC scales' reliability and validity. Additionally, this study evaluated if including additional measurement methods contributed to improved predictions over SR methods alone. Overall, the findings offer promising evidence for the scale's validity while also highlighting improvement opportunities.

Likert Items

Currently, SR Likert scales are the most popular method used in personality assessment, as they often provide valid and reliable measurements. Consistent with previously developed scales, the ACT® Tessera® SR Likert scales provided evidence for validity and reliability in study 1. The ACT® Tessera® SR Likert scales demonstrated acceptable reliability for each social-emotional trait measured. However, a CFA testing the SR Likert scales' internal structure validity yielded conflicting fit indices, making it challenging to confidently discern the scale's ability to reflect the FFM validly. While the CFI and TLI challenged ACT® Tessera®'s applicability to the FFM, these findings are not entirely surprising when considered alongside the broader literature. Previous studies validating several FFM-aligned scales also struggled to obtain acceptable model fit using CFA (Hopwood & Donnellan, 2010). Therefore, challenges with internal structure may not be as problematic as they initially appear.

As previously stated, difficulty obtaining an acceptable model fit for FFM-based scales is common when using CFA. However, several of the scales that have struggled to obtain an acceptable model fit are still widely accepted because they demonstrate sound criterion validity, and more recent research has provided internal structure validity evidence when using exploratory factor analytic techniques (Borkenau & Ostendorf, 1990; Donnellan, et al., 2006; Hopwood & Donnellan, 2010). Therefore, Hopwood and Donnellan (2010) argue that these challenges reflect problems with using CFA to evaluate FFM-based measures internal structure, rather than an issue with the FFM itself.

Social-emotional traits are complex to conceptualize, and there may be several reasons for poorly fitting models, such as cross-loading factors and correlated residuals

within the same trait items (Hopwood & Donnellan, 2010). In response to these challenges, Hopwood and Donnellan (2010) suggest that future research explore alternatives to using CFA to validate FFM scales, such as exploratory factor analysis. Simultaneously, until better evaluation methods are available, research suggests that researchers consider such scales more holistically (i.e., construct, criterion validity) instead of relying on traditional cut-offs to evaluate model fit.

Analyzing the SR Likert CFA factor loadings to examine the scales' internal structure validity further is promising. All but three items loaded satisfactorily on their intended factor, with only two Resilience and one Leadership item yielding loadings slightly below the desired level ($<.50$). The weak Resilience items stated: "I cope well with last-minute changes to assignments," and "I am comfortable changing direction in the middle of a class project." When comparing these weaker Resilience items to those that loaded better on the factor, they reflected an individual's internal experiences rather than outward behavior. The other items that loaded well on the Resilience scale illustrated *behavioral responses* to unfavorable situations, such as, "I speak calmly even when I am angry," and "I stay calm during disagreements." These stand in contrast to the internal emotional responses reflected in the "weaker items." Therefore, these items may load better onto the Resilience factor if rewritten to reflect behavioral responses to the described situations.

The SR Likert item that yielded the weakest factor loading was an item that read, "I am often able to convince others to agree with me." Intuitively this item appears to reflect the intended Leadership factor, defined as an individual's ability to "demonstrate assertiveness, influence, optimism, and enthusiasm" (ACT®, 2018). It taps explicitly

into an individual's assertiveness and influence. However, to capture the essence of influence, this item considers another person's response to the respondent's behavior. This overlap with another individual's behavior may confound the measure because the rater is required to consider both their likelihood to behave in a specific situation and *another person's* likelihood of behaving in a specified manner. For example, if the respondent surrounds himself or herself with people who are low on the Curiosity trait, then they may be unlikely to convince them to agree with them regardless of how high they are on Leadership. Rewriting this item to focus solely on the respondent's behavior may yield factor loadings that load more consistently with other items in this domain.

Situational Judgment Tests

Study 1 also assessed the SJT scale reliability. According to traditional measurement conventions, Grit and Leadership demonstrated "acceptable" reliability, whereas the Teamwork, Resilience, and Curiosity scales demonstrated "questionable" reliability (George & Mallery, 2003). However, conventional standards for evaluating SJT scale reliability are limited because SJTs are inherently multi-dimensional. SJTs present an individual with a specific situation and ask them to indicate the likelihood of responding in specified ways. These real-life situations require a response that may combine the expression of several constructs, which confounds measures to some degree (Corstjens, Livens, & Krumm, 2017). When comparing the present findings with the internal consistencies reported by SJT methods in the literature, Cronbach's alpha ratings outperform expectations, with the average alpha documented in the literature as .57 (Campion, Ployhart, MacKensie, 2014). Therefore, the ACT® Tessera® SJT scales provide measures that are sufficiently reliable when measuring social-emotional skills.

Similar to the SR Likert scales internal structure analyses, variable fit indices also characterized the SJTs internal structure analyses. Considering the previously discussed model fit challenges using CFA and challenges with SJT multidimensionality, these findings are unsurprising. However, when reviewing the scale's factor loadings, some scales performed better than others. The Leadership and Grit items generally demonstrated satisfactory loadings on their intended factor, with only one item per scale loading below .50. However, the Teamwork, Curiosity, and Resilience loadings were more concerning. Therefore, the following discussion reviews these weaker items and makes data-based recommendations.

Problems with the SJT Curiosity and Resilience internal structure were subtle and appeared to be caused by the SJTs multidimensional nature. While the items that loaded poorly appeared to primarily represent their intended trait, these traits also appeared to reflect the Grit scale secondarily. This observation is consistent with research that has demonstrated individuals do not express traits in isolation in applied situations and that there is an interaction between traits (Merz & Roesch, 2011). Therefore, the internal structure of these scales may not be as poor as it initially appears.

Reviewing the SJT Teamwork factor loadings reveals that responses to one of the two SJT situations yielded more substantial loadings on the intended factor than the other. Responses to the items that followed the second SJT Teamwork situation yielded loadings that ranged from .22 to .43. The weaker items asked the individual to identify the likelihood that they would respond in specified ways to a situation in which they offended others during a disagreement. While remedial actions in this situation certainly would reflect Teamwork, it may be unlikely that someone high on this trait would find

themselves in this situation in the first place. Therefore, it may be difficult for individuals who demonstrate the highest levels of this trait to rate their likelihood of displaying each response. If future scale iterations revise these items to reflect the actions that someone may take during a disagreement to *avoid offending* others, the item loadings and overall internal structure may be improved.

Forced Choice

Regarding the FC scales' reliability and validity, this study faced substantial limitations to studying these scales. While the data reflecting the scales' reliability is questionable, these analyses were limited by the ipsative scoring procedure used (See Limitations for more detail). However, the FC scales generally exhibited convergence with SR and SJT scales measuring the same constructs. Additionally, the FC items also demonstrated similar relationships with GPA when compared the FC and SJT items. Therefore, while the FC analyses are insufficient to confidently conclude the FC scales' reliability and validity, there is some indication that future analyses may support the scales validity.

Convergent and Discriminant Validity

As previously discussed, convergent and discriminant validity analyses scrutinized correlations between different item types measuring the same construct. Overall, correlations between the SR Likert and SJT scales demonstrated good convergent validity evidence for all five traits measured by ACT® Tessera®. The same was generally true for FC correlations with SR Likert scales. In terms of the correlations between FC and SJT scales, the Grit, Teamwork, Curiosity, and Resilience scales demonstrated convergent validity. However, it is essential to note that the convergence

between FC and SJT Teamwork and Curiosity scales was low quantitatively. Additionally, the SJT and FC Leadership scales demonstrated poor convergent validity, as the correlation between the items was weaker than the correlation between SJT Leadership and FC Grit scale. Notably, this failure to converge appears to be a function of the SJT Leadership scale's difficulty with discriminant validity discussed below.

While the ACT® Tessera® scales generally demonstrated good discriminant validity, the SJT Leadership scale struggled to distinguish itself from different FFM traits. Specifically, the SJT Leadership scale correlated unexpectedly high with multiple other FC and SR Likert scales intending to measure different traits, especially Grit. Therefore, measures provided by the SJT Leadership scale may not truly be reflecting the intended construct. This finding was unexpected given the SJT Leadership scales' high factor loadings demonstrated by the CFA.

Upon closer review, the SJT Leadership scale's difficulty with discrimination appears to stem from ACT® Tessera®'s focus on school-related situations. The SJT Leadership measures appear to measure the intended trait primarily. However, these items place the student in a situation in which they would need to use their Leadership trait to achieve academically, which also reflects Grit. This again highlights some of the challenges with multidimensionality when using SJTs. However, considering these findings alongside the predictive validity data discussed below, this scale should be revised to provide purer Leadership measures.

Predictive Validity

ACT® Tessera® demonstrated some indication of predictive validity.

Specifically, all Conscientiousness scales demonstrated significant positive correlations

with college GPA. These results are consistent with findings reported in the empirical literature that has repeatedly reported a significant relationship between this trait and achievement (Busato, et al., 1999; Chamorro-Premuzic & Furnham, 2003; Conard, 2006; Paunonen & Ashton, 2001). However, contrary to my hypotheses, SJT Leadership was associated with college GPA. The significant correlation between the SJT Leadership scale and college GPA provides further evidence that this scale may be reflecting Grit more than the intended trait. This is especially true when compared to the correlations between the other Leadership scales (SR Likert and FC), which yielded negligible correlations between Leadership and GPA. These findings bolster the argument that the SJT Leadership scale should be modified to provide a purer Leadership measure.

Notably, the SR Teamwork scale also significantly correlated with college GPA, whereas the SJT and FC and Teamwork scales demonstrated a non-significant relationship with GPA. The difference in these scales' relationship with college GPA was surprising given the convergent validity demonstrated between these scales. However, the magnitude of the correlation between SR Teamwork was small. Item analysis detected no glaring concerns with the SR Teamwork scale, and therefore, no recommendations for revision are made.

Incremental Validity

Another meaningful finding established by study 1 is that despite predictions that multiple response methods should improve predictive validity, this study's results do not fully support this notion. While SJT's added some incremental validity in predicting college GPA above and beyond SR Likert scales, the amount of additional variance accounted for by the SJT scales was meager. Additionally, adding FC measures into the

prediction offered no significant improvement. Therefore, using multiple measurement methods may not be advantageous over traditional SR Likert methods as hypothesized.

General Conclusions and Considerations

Overall, ACT® Tessera® exhibits strengths in measuring social-emotional learning skills. Notably, the ACT® Tessera® SR scales generally provided acceptable reliability and validity for assessing social-emotional skills in a university population. While the preceding discussion identifies several items that could be revised to improve the SR scales internal structure validity, statistical analyses generally provide preliminary evidence to support the scale's utility for measuring FFM-based social-emotional learning traits. Additionally, Grit performed well on most reliability and validity analyses. ACT® Tessera® reliably measured, demonstrated sound convergent/discriminant validity, and evidenced strong predicative validity when considering Grit.

Limitations and Directions for Future Research

One limitation of this study is the ipsative scoring approach used to score the FC scales. Using ipsative data is problematic because it violates Classical Test Theory assumptions. Therefore, reliability estimates provided by Cronbach's alpha are likely distorted (Meade, 2004). While Item Response Theory (IRT)-based scoring methods are available to overcome the limitations of ipsative data, calculating these scores was beyond the scope of this paper. However, future studies should evaluate the reliability of these scales using the IRT approach. Doing so can enhance the understanding of scale's reliability and further identify areas for improvement.

Another limitation was imposed by using CFA to evaluate the SR and SJT scales internal consistency. As previously mentioned, challenges with using CFA for scales reflecting the FFM are documented in the literature (Hopwood & Donnellan, 2010). Therefore, this study was limited in its ability to qualitatively substantiate ACT® Tessera®'s internal structure validity. Currently, in the literature there are some developing statistical methods that may be used to test internal structure in the future, which include using Exploratory Factor Analysis (Hopwood & Donnellan, 2010). However, at this time this method is not fully developed. Over time, as more substantial methods are supported for evaluating FFM internal structure, future studies should follow this line of research.

Additional limitations for this study surround the representativeness of the sample. The sample utilized in study 1 reflects the population at one metropolitan university. While the sample was moderately sized and relatively diverse, the data was derived from only one university. The sample at St. John's University is unique as it is located in an urban borough of New York City. However, this sample cannot reliably represent the broader university student population across the country. Therefore, before this measure is confidently used to assess social-emotional traits in different populations, data regarding its validity should be collected and compared across several additional samples.

CHAPTER 2

Study 2 sought to investigate the predictive power of ACT® Tessera®'s social-emotional traits above and beyond commonly used university achievement predictors, such as standardized test scores, intelligence, and high school Grade Point Average (GPA). Study two combined study 1 data, with intelligence data provided by a subset of participants. Data were evaluated through hierarchical regression analyses to assess if personality predicted university achievement while controlling for traditional predictors.

Hypotheses

Given findings reported in the literature that suggest that social-emotional traits add to university achievement predictions beyond intelligence (Bratko, Chamorro-Premuzic, & Sacks, 2006; Nofle & Robins, 2007; Ziegler, et al., 2010), I expected ACT® Tessera®'s social-emotional measures to predict university achievement above and beyond traditionally used achievement predictors. Notably, I expected that social-emotional traits would uniquely contribute to university achievement beyond both SAT® scores and intelligence, as these predictors are highly correlated (Frey & Detterman, 2004).

Method

Participants and Procedures

In order to recruit study 2 participants, all study 1 participants were contacted through their university email addresses ($N=706$). Participants were only eligible to participate if they participated in study 1. Potential study participants were contacted with an email alert that explained the study expectations and provided a link to the IRB-approved consent form on the Qualtrics survey platform (Appendices E & F). After a

student provided informed consent, they received an individualized assessment link via email. The assessment link brought participants to a cognitive ability assessment, Mindprint Learning® (see description below). Participants who did not provide consent, were sent additional reminders of the opportunity to participate 3, 5, 7, and 10 days after the initial email was sent (Appendix G). As an incentive to participate in the study, all participants were allowed to request their survey results. After the survey closed, the participants' survey results were combined with the first study's measures. Specifically, college GPA, SAT®-Verbal, SAT®-Math, and ACT® Tessaera® scores were used in conjunction with the Mindprint Learning® cognitive measures for the present study.

61 students ($N=62$) completed the Mindprint Learning® assessment, and their performance was retained for analysis. However, it is important to note that these data were only used for regression 1. Participants ranged from age 18 to 22, with a mean age of 19.70 ($SD=1.160$). 7 freshmen (12.7%), 14 sophomores (24.5%), 23 juniors (41.8%), and 10 seniors (18.2%) participated in study 2. 29.1% percent of participants were honor students. The sample was 72.7% female. Participants identified as White (36.4%), Black/African American (21.8%), Asian (20.0%), Hispanic (10.9%). 5.5% identified with 2 or more races, and 3.6% indicated an unknown ethnicity. Demographic characteristics for study two participants appear in Table 8. Data for regressions 2-4 were obtained from the final study 1 sample (see demographic characteristics reported above).

Table 8*Demographic Characteristics (Study 2)*

	N	Percentage (%)
Gender		
Female	40	72.7
Male	14	25.5
Age		
18	10	18.2
19	12	21.8
20	20	36.4
21	8	14.5
22	4	7.3
Class Standing		
Freshman (1 st year)	7	12.7
Sophomore (2 nd year)	14	25.5
Junior (3 rd year)	23	41.8
Senior (4 th year)	10	18.2
Ethnicity		
2 or more races	3	5.5
Asian	11	20
Black or African American	12	21.8
Hispanic	6	10.9
White	30	36
Unknown	2	3.6
Honors Status		
Honors Student	16	29.1
Non-Honors Student	39	70.9

Measure: Mindprint Learning®

Mindprint Learning®, also known as the Penn Computerized Neurocognitive Battery, is a measure that the Brain Behavior Lab developed at University of Pennsylvania's Perelman School of Medicine (Moore et al., 2015). Previous studies have reported on the measure's reliability, construct validity, and internal structure validity (reliability estimates accuracy score $\alpha=.55-.95$; reliability estimates accuracy score $\alpha=.77-.97$) (Gur, et al., 2010; Gur, et al. 2012; Moore, et al., 2015). The measure was designed to assess accuracy and speed in specific neurobehavioral domains using tests that were previously validated with functional neuroimaging. The neurobehavioral domains measured, include: Attention, Working Memory, Visual Memory, Processing

Speed, Verbal Memory, Visual Motor Speed, Flexible Thinking, Verbal Reasoning, Abstract Reasoning, and Spatial Perception. Mindprint Learning® is a computer administered cognitive battery that takes approximately one hour to complete. Mindprint Learning® provides z-scores to represent an individual's performance in an assessed domain. Higher scores on these measures indicate better performance.

Analyses

To evaluate this study's hypotheses, I used SPSS Statistics 26 (IBM, 2019) to conduct all statistical analyses. First, I conducted a hierarchical regression to evaluate social-emotional skills' predictive ability when controlling for intelligence and standardized test scores. In order to determine the order that predictors were entered into the regression equation, the empirical literature was consulted. Previous research has established a significant positive correlation between general intellectual ability and standardized exam scores (Frey & Detterman, 2004; Koenig, Frey, & Detterman, 2008; Syverson, 2007). Both of these predictors have also demonstrated significant positive relationships with college academic performance. This relationship is so well established that some researchers have even categorized the SATs as an intelligence test (Nofle & Robins, 2007). On the other hand, others have postulated that the SAT® considers additional factors that may account for variance in GPA (Coyle & Pillow, 2008). Therefore, when evaluating college GPA predictors, cognitive ability was entered into the hierarchical regression first, then standardized test scores were entered second. Additionally, social-emotional traits were entered third to determine if they offer predictive validity beyond traditional predictors (i.e., intelligence, standardized exams).

Additional hierarchical regression analyses were conducted to evaluate the predictive relationship between standardized test scores and social emotional traits in relation to college GPA when utilizing a larger sample. Finally, in line with this study's objective of determining if social-emotional traits can contribute to academic success predictions in the college admissions process, a final hierarchical regression considered both standardized tests, and social emotional traits' ability to predict college GPA when controlling for high school GPA.

Results

Regression 1

A three-stage hierarchical multiple regression was conducted with college GPA as the dependent variable. After cases were deleted listwise, 38 cases were included in the analysis. Intelligence variables, including, Attention, Working Memory, Visual Memory, Processing Speed, Verbal Memory, Visual Motor Speed, Flexible Thinking, Verbal Reasoning, Abstract Reasoning, and Spatial Perception, were entered into block one. SAT® scores including, SAT®-Verbal, and SAT®-Mathematics scores, were entered into block two. ACT® Tessera®'s social-emotional traits, including Grit, Teamwork, Resilience, Curiosity, and Leadership, were entered at stage three. Figure 9 shows model summary statistics and the change in R^2 .

Table 9*Hierarchical Regression: Intelligence, Standardized Tests, Social-Emotional Traits**(n=38)*

	R	R²	Adjusted R²	SE	R² Change
Intelligence	.47	.22	.02	.29	--
Intelligence, Standardized Test	.63	.39	.20	.56	.17*
Social-Emotional Traits	.72	.52	.25	.56	.13

*Note. *p=.01***Regression 2**

A two-stage hierarchical multiple regression was conducted with college GPA as the dependent variable. After cases were deleted listwise, 443 cases were included in the analysis. SAT®-Verbal, and SAT®-Mathematics scores were entered block one. The social-emotional traits (*Grit, Teamwork, Resilience, Curiosity, and Leadership*) were entered at stage two. Table 10 shows model summary statistics and the change in R².

Table 10*Hierarchical Regression: Standardized Tests, Social-Emotional Traits (n=433)*

	R	R²	Adjusted R²	SE	R² Change
Standardized Test	.27	.08	.08	.46	--
Standardized Test, Social-Emotional Traits	.50	.25	.24	.42	.17*

*Note. *p<.01***Regression 3**

A two-stage hierarchical multiple regression was conducted with college GPA as the dependent variable. After cases were deleted listwise, 443 cases were included in the

analysis. The ACT® Tessera® social-emotional traits (*Grit, Teamwork, Resilience, Curiosity, and Leadership*) scores were entered block one. SAT®-Verbal and SAT®-Mathematics were entered at stage two. Table 11 shows model summary statistics and the change in R².

Table 11

Hierarchical Regression: Social-Emotional Traits, Standardized Tests (n=433)

	R	R ²	Adjusted R ²	SE	R ² Change
Social Emotional Traits	.39	.15	.14	.44	--
Social-Emotional Traits, Standardized Test	.50	.25	.24	.42	.10*

Note. * $p < .01$

Regression 4

A three-stage hierarchical multiple regression was conducted with college GPA as the dependent variable. After cases were deleted listwise, 444 cases were included in the analysis. High-school GPA was entered into block one. SAT® scores including, SAT®-Verbal, and SAT®-Mathematics scores, were entered into block two. ACT® Tessera®'s social-emotional traits, including Grit, Teamwork, Resilience, Curiosity, and Leadership, were entered at stage three. Table 12 shows model summary statistics and the change in R².

Table 12*Hierarchical Regression: College GPA, Standardized Exams, Social-Emotional Traits*

	R	R²	Adjusted R²	SE	R² Change
High School GPA	.47	.22	.22	.42	--
High School GPA, Standardized Test	.48	.23	.22	.43	.00*
High School GPA, Standardized Test, Social-Emotional	.57	.32	.31	.40	.09*

*Note. *p<.01***Discussion**

Study two evaluated social-emotional traits' ability to predict university academic performance beyond traditional achievement predictors. This study utilized a series of hierarchical regression analyses to explore the predictive relationship between social-emotional traits, intelligence, high school GPA, and standardized testing when predicting college achievement.

Regression one found that collectively, intellectual abilities measured by Mindprint Learning® did not account for a significant proportion of college GPA variance. However, when standardized test scores entered the equation, they significantly improved the predictive model. These findings were somewhat unexpected given empirical evidence that the SATs are highly intelligence-loaded (Frey & Detterman, 2004). However, the increase in predictability when adding standardized tests into the model is consistent with research that documented that SATs offer a unique contribution to college GPA prediction beyond general intelligence (Coyle & Pillow, 2008). Another surprising finding demonstrated by this analysis was that contrary to this study's hypothesis, social-emotional traits did not account for a significant proportion of college

GPA variance beyond cognitive ability and standardized test scores. Again, these findings were surprising given studies documented in the literature that suggests that personality may predict academic performance more strongly than intelligence (Conard, 2006; Furnham & Chamorro-Premuzic, 2004; Furnham, Chamorro-Premuzic, & Mcdougall, 2003).

The previously discussed findings that contradicted reports in the empirical literature highlighted limitations posed by the sample. The small student sample utilized for regression one may have obscured the analysis. While diligent efforts were made to obtain the largest sample possible, only 38 complete data sets were available for the analysis when deleting cases listwise. Therefore, the difference between the present study's findings relative to reports in the literature is likely a result of insufficient power, rather than a meaningful discrepancy between findings (Green, 1991).

Given the first analysis's limitations, regressions two and three evaluated the ACT® Tessera® traits' predictive value using more data. Collectively, regression two and three indicated that both social and emotional traits and standardized exam performance are useful predictors of college GPA. Specifically, each variable offered a unique contribution to the prediction when controlling for the other. Therefore, these findings support the notion that college admissions committees should consider social emotional traits in their selection processes. While these findings may have significant implications for practice, they are limited because they only consider two achievement predictors whereas, college admissions committees often have additional data available such as high school GPA. Therefore, regression four adds significant value to this discussion.

Regression four found that high school GPA significantly accounted for 47% of the variance in college GPA. However, standardized test scores yielded only a small increase in the prediction beyond high school GPA. Additionally, consistent with this study's hypothesis, social-emotional skills contributed to a significant increase in the college GPA prediction. Together these findings suggest that utilizing standardized test scores to predict an individual's potential to succeed in college may offer little benefit. Instead, it may be more useful to evaluate high school GPA in conjunction with social-emotional measures when predicting college GPA.

Regression four's findings lend support to the argument that colleges and universities should adopt test-optional policies. Given the minimal increase that standardized tests added to the college GPA prediction beyond high school GPA in the present sample, these exams' costs may not be worth the benefit. Families make considerable monetary investments in standardized testing when a student wishes to attend college. Some costs come directly from test developers when students pay a fee to take the exam (many students take the exam multiple times) and send their scores to schools. Students also incur secondary costs by purchasing exam preparation materials. Numerous test preparation materials are available to students, including books, online programs, courses, and online tutoring, which may cost a family up to hundreds or thousands of dollars (Robinson, 2019). When considering these costs, it is hard to justify the investment when college admissions committees may make successful predictions without these scores.

While regression four's findings may have significant implications, they must be evaluated alongside the broader literature. The findings reported by regression four are

inconsistent with The College Board®'s reported findings that the SATs contributed to a 15 percent increase in college GPA above and beyond high school GPA (Westrick, et. al., 2019). However, it is essential to note that The College Board®'s analyses relied on self-reported high school GPA data and only collected first-year college GPA data. Utilizing *self-reported* high school GPA is problematic because it may have increased the likelihood of measurement error. For example, an individual may under or over report high school GPA for various reasons, including concerns with social desirability, carelessness, or forgetfulness. Therefore, the self-reported measures may have obscured the analyses provided by The College Board®.

Another flaw in The College Board®'s methodology surrounds the fact that the researchers relied solely on the first-year GPA. Focusing analyses on first-year GPA measures may also be problematic because students take foundational classes their first year, and classes tend to increase in difficulty over college levels. Therefore, these measures may not be truly representative of a student's overall college GPA. Overall, these confounds present in The College Board®'s study may have contributed to the discrepancy with the present study's findings.

Limitations and Future Directions

It is essential to consider this study's findings with some potential limitations. As previously alluded to, regression one was hindered by the available sample because it did not provide sufficient power to obtain a significant effect. The small sample size was likely due to a flaw in the study's design that delayed the participants' access to the measure at the time of consent. Additionally, it is unlikely that a sample of 38 students is representative of the overall college population. These sample limitations interfered with

this study's ability to understand the true predictive relationship between cognitive ability, standardized tests, and social, emotional traits. Given the documented relationship between intelligence and standardized tests (Frey & Detterman, 2004; Koenig, Frey, & Detterman, 2008; Syverson, 2007), future researchers should evaluate this relationship more thoroughly, with a larger sample.

Additional limitations surround this study's methodology. Despite the small sample used, there are additional concerns with using an intelligence test that was administered remotely. Traditional intelligence measures are administered under strictly standardized conditions. By allowing participants to complete this measure remotely, the study was vulnerable to confounding factors. For example, participants may have completed the assessment in a noisy, and distracting environment, or they may have had someone else complete the assessment for them. Therefore, future research should also evaluate the relationship between intelligence and college performance, using traditional assessment measures.

Another limitation facing this study is the fact that the SATs were used to represent "standardized exams" in all analyses. Unfortunately, among the data provided by the Office of Institutional Research, there were few ACT® scores provided because these scores are not required for admission to St. John's University. While there is data to suggest that the ACT® and SATs are both highly correlated with GPA (SAT® $r=.72-.86$; ACT® $r=.61-.7$) (Syverson, 2007; Frey & Detterman, 2004; Koenig, Frey, & Detterman, 2008), the findings of this study should be replicated with ACT® scores to increase the generalizability of these findings to other standardized admissions exams.

Implications for School Psychologists

Improving Applied Practice

The present study has several implications for school psychologists. Primarily, according to the National Association of School Psychologists (NASP), “school psychologists are qualified members of school teams that support students’ ability to learn... [by applying] expertise in mental health, learning, and behavior.” Among school psychologists’ many duties, they help schools improve academic achievement by conducting assessments and making appropriate recommendations (“Who Are School Psychologists?,” n.d.). However, school-based assessments typically rely on intelligence measures and fail to account for social-emotional traits’ impact on success. Considering the evidence supporting FFM based traits’ associations with academic achievement (Poropat, 2009), assessments conducted including social-emotional learning measures would likely improve student success predictions. However, to date, there are few of these measures available to be used in an academic setting. The preliminary evidence demonstrated in study one regarding ACT® Tessera®’s reliability and validity is exciting for the field of school psychology. If future scale iterations provide more evidence to support ACT® Tessera®’s validity, school psychologists may introduce this new tool to their repertoire soon.

Advocacy

Additionally, this study’s findings present an opportunity for school psychologists, especially those employed in university settings, to advocate for better university admissions procedures on their students’ behalf. “School psychologists strive to ensure that all [students] have equal opportunity to participate and benefit from school

programs... school psychologists take steps to foster a school climate that is supportive, inclusive, safe, accepting, and respectful toward all persons..." (The Professional Standards of the National Association of School Psychologists, 2020, p.44). In light of the arguments outlined in the literature review that highlight the limitations of standardized testing on students from marginalized populations and the present study's finding that standardized tests do not significantly contribute to academic success predictions beyond high school GPA has significant implications.

In addition to encouraging diversity, school psychologists should advocate for changes in admissions practices given their responsibility to promote systems change to benefit all students and clients (The Professional Standards of the National Association of School Psychologists, 2020, p.54). Considering this study's findings with the previously mentioned costs associated with standardized exams, these findings again open up the door to advocacy. More specifically, it would benefit all students to abandon standardized exams if they do not offer to provide admissions committees with meaningful data for predicting success. In particular, if future studies confirm this dissertation's findings that suggest that standardized tests offer little value in college prediction, school psychologists should be among the first professionals to advocate a change in college admissions to promote diversity and to benefit all students.

Graduate Admissions

In addition to advocacy opportunities, this dissertation's findings also offer preliminary evidence to improve the graduate school admissions process. Like admission to undergraduate programs, school psychology training programs often require prospective students to submit standardized test scores to accompany their other

application data (i.e., Graduate Record Exam). Therefore, if future research expands upon this study's findings to a graduate school population, the school psychology admissions process may be improved. For example, suppose future research establishes a predictive relationship between social-emotional traits and graduate success outcomes beyond the GRE (i.e., GPA, job performance, etc.). In that case, this may help programs select the highest quality applicants for the program and, thus, the field of school psychology.

APPENDICES

Appendix A

Initial Participant Contact



Good Morning [Student]!

You have been provided the opportunity to participate in a study that involves St. John's University undergraduate students. If you agree to participate in this study, you will be involved in a study that is working to understand the factors that contribute to academic success. Additionally, this research will help identify the ways in which students can be provided the opportunity to work to their potential.

Please be aware that you will be receiving a formal e-mail invitation later on today.

Best Regards,
Allison Murray, M.S.
Department of School Psychology

Marlene Sotelo-Dynega, PsyD, ABSNP, NCSP
Associate Professor
Director, Graduate Programs in School Psychology

Appendix B

Study 1 Formal Invitation to Participate



Dear Student:

I am a graduate student in the Psychology Department here at St. John's University. I will be conducting a research project titled "The Effects of Personality on Academic Success Among College Students". This study will evaluate how certain factors relate to academic success among undergraduate students like yourself.

Your participation in this study is confidential and strictly voluntary. If you choose to participate and then change your mind, you will be able to withdraw at any time. You will electronically presented with a consent form that will detail your rights as a participant in this study. Once you have provided consent, you will be directed to a number of rating scales. The completion of this questionnaire should take approximately x minutes to complete. To participate, please click here [insert hyperlink] to be directed to the online survey.

Some things you should know before beginning:

- If you chose to participate, you will be asked to sign the consent form using your St. John's University X-Number.
- We will be using your X-Number to pull pertinent information (i.e. high school grade point average, Scholastic Aptitude Test [SAT] scores, socioeconomic status, etc.) to help us determine the best predictors of academic success.
- We may contact you to invite you to second phase of this study where we will conduct a brief assesment of your cognitive functioning.
- To monitor academic success we will be accessing your current grades and enrollment status at St. John's University.

If you have any questions about this project or your participation in it, you may email Allison Murray at allison.murray16@stjohns.edu or Dr. Marlene Sotelo at sotelodm@stjohns.edu.

Thank you for your participation and assistance!

Sincerely,

Allison Murray, M.S.
Department of Psychology

Marlene Sotelo-Dynega, PsyD, ABSNP, NCSP
Associate Professor
Director, Graduate Programs in School Psychology

Appendix C

Study 1 Informed Consent Agreement



I hereby agree to serve as a participant in the research project titled “The Effects of Personality on Academic Success Among College Students.” It has been explained to me that the purpose of this study is to examine the effects of personality on academic success. By signing this Informed Consent Agreement, I am acknowledging that I understand and agree to the following:

- I am at least 18 years of age.
- I am a undergraduate student at St. John’s University.
- I am aware that the online assessment will take approximately 25 minutes to complete. Additionally, I understand that if I am randomly selected to participate in an additional cognitive assessment, I will be contacted by the researcher again.
- I am aware that if I am asked to participate in a cognitive assessment, my participation will take approximately 30 minutes and that I will be asked to meet with the researcher in person.
- I understand that the data collected during the assessment will be remain confidential and will be stored in a locked, secure place that only the investigators will have access to.
- I have been informed that my St. John’s University Identification X-Number will be used to access information about me that was submitted by me (or on my behalf) to the University; I understand that this information includes: age, total credits earned, class year at SJU (i.e. freshman, sophomore, junior, senior), academic term admitted, high school Grade Point Average, Scholastic Aptitude Test (SAT) scores, ACT scores, major status, major, socio-economic status, first generation of college attendance (i.e., first person in family to go to college), number of delinquencies prior to the beginning of the fall 2018 semester, gender, ethnicity, and current college GPA.
- I understand that the data collected by the investigators of this study will be shared with ACT, Inc. for research purposes.
- I understand that the risks to me, if any, are minimal, but if I feel distressed by any of the questions or procedures at any point, I will contact the St. John’s University Center for Psychological Services at 718-990-1900.
- I understand that if I have any questions about this project or my participation in it,

I can call Marlene Sotelo-Dynega at 718-990-1545 or email her at sotelodm@stjohns.edu. For questions about my rights as a research participant, I may also contact Raymond DiGiuseppe, PhD at the university's Human Subjects Review Board at digiuser@stjohns.edu.

- I understand that my participation in this project is completely voluntary, and that my choice of whether to participate in this project will not jeopardize my relationship with St. John's University.
- I understand that my participation in this study is voluntary, and that I have the right to withdraw at any point before or during the study.
- I understand that as an incentive for participating in this study, I may contact the researcher to obtain my results of the personality assessment. I understand that I must reach out to the researcher via email at allison.murray16@stjohn.edu to obtain this benefit. Further, I understand that it may take a few weeks to get my results and that they will be emailed to me.

By entering my X number below, this will serve as my electronic signature. By doing so I attest that I have read and agree to the above statements.

X-Number

Date

Appendix D

Study 1 Reminder Email



Dear [Student]:

You received a formal invitation to participate in a study [x] days ago, yet you have yet to respond. If you are interested in participating, please follow the link below to the survey. If you are not interested, I thank you for your cooperation.

[Insert Hyperlink]

Wishing you all the best,

Allison Murray, M.S.,
Department of Psychology

Marlene Sotelo-Dynega, PsyD, ABSNP, NCSP
Associate Professor
Director, Graduate Programs in School Psychology

Appendix E

Study 2 Invitation to Participate



Good Morning!

Thank you for your completing phase one of our study and assisting us with our research thus far. Your participation is contributing to our understanding of academic success indicators.

This email serves to provide you with the opportunity to participate in the second phase of this study. If you agree to participate, you will take part in a cognitive assessment to supplement the information you previously provided. Again, your participation in this study is confidential and strictly voluntary. If you choose to participate and then change your mind, you will be able to withdraw at any time. Your participation in the cognitive assessment should approximately 60 minutes. The assessment and will require you to answer some questions and complete some activities.

If you wish to participate, please click the link below to sign the informed consent form. After you click the link, you can expect to receive a follow up email with a link to complete the assessment.

[link](#)

Please do not hesitate to contact us with any questions, comments or concerns.

Best Regards,
Allison Murray, M.S.
Student Researcher
Department of Psychology

Marlene Sotelo-Dynega, PsyD, ABSNP, NCSP
Associate Professor
Director, Graduate Programs in School Psychology

Appendix F

Study 2 Informed Consent



I hereby agree to serve as a participant in phase two of the research project titled “The Effects of Personality on Academic Success Among College Students.”

It has been explained to me that the purpose of this study is to examine the predictors of academic success. By signing this Informed Consent Agreement, I am acknowledging that I understand and agree to the following:

- I am at least 18 years of age.
- As a participant in phase two of this study, I agree to take part in a web-based assessment of cognitive abilities. I am aware that my participation will take approximately 1 hour.
- I understand that the data collected during the cognitive assessment will be remain confidential and will be stored in password protected files, only accessible to the researchers.
- I have been informed that by consenting to participate, the researcher will obtain information about me that was submitted by me (or on my behalf) to the University; I understand that this information includes: age, total credits earned, class year at SJU (i.e. freshman, sophomore, junior, senior), academic term admitted, high school Grade Point Average, Scholastic Aptitude Test (SAT) scores, ACT scores, major status, major, socio-economic status, first generation of college attendance (i.e., first person in family to go to college), number of delinquencies prior to the beginning of the fall 2019 semester, gender, ethnicity, and current college GPA.
- I understand that all data collected by the investigators of this study will be shared with ACT, Inc. for research purposes.
- I understand that the results of my Mindprint Learning (cognitive) assessment, my ACT score, my SAT score, my major status, my high school GPA, and my college GPA will be shared with Mindprint Learning, for research purposes.
- I understand that by consenting to participate in phase 2 of this study, the results of the personality assessment I completed during phase 1 will be shared with Mindprint Learning.
- I understand that by consenting to participate in phase 2 of this study, the results of my Mindprint Learning Assessment (phase 2) will be shared with ACT, Inc.
- I understand that the investigators of this study will access my grades and enrollment status during my time at St. John’s University. I understand that any information gathered about me and my functioning

will be maintained as confidential and will only be used for the purposes of this study.

- I understand that the risks to me, if any, are minimal, but if I feel distressed by any of the questions or procedures at any point, I will contact the St. John's University Center for Psychological Services at 718-990-1900.
- I understand that if I have any questions about this project or my participation in it, I can call Marlene Sotelo-Dynega at 718-990-1545 or email her at sotelodm@stjohns.edu. For questions about my rights as a research participant, I may also contact Raymond DiGiuseppe, PhD at the university's Human Subjects Review Board at digiuser@stjohns.edu.
- I understand that my participation in this project is completely voluntary, and that my choice of whether to participate in this project will not jeopardize my relationship with St. John's University.
- I understand that my participation in this study is voluntary, and that I have the right to withdraw at any point before or during the study.
- I understand that as an incentive to participate in this study, I may contact the principle investigator by May 13, 2020 to receive the results of my Mindprint Learning Assessment and that it may take several weeks to receive the results of this study.

By entering my X-number below, I attest that I have read and agree to the above statements.

X-Number

Appendix G

Study 2 Reminder Email



Study 2 Reminder Email

Good Afternoon!

You received a formal invitation to participate in a study approximately one week ago and have not yet responded. This survey will close on August 12, 2020. If you are interested in participating, please follow the link below to consent to participate in the survey. If you are not interested, I thank you for your consideration.

https://stjohns.az1.qualtrics.com/jfe/form/SV_3EgsOYX60em8F3D

Wishing you all the best,

Allison Murray, M.S
Student Researcher
Department of Psychology

Marlene Sotelo-Dynega, PsyD, ABSNP, NCSP
Associate Professor
Director, Graduate Programs in School Psychology

REFERENCES

- ACT®. (2020, October 14). *U.S. high school graduates, underserved students will face significant challenges in college according to new ACT® achievement data.*
- ACT® Newsroom and Blog. <https://leadershipblog.ACT®.org/2020/10/ACT®-grad-class-2020-achievement-data.html#:~:text=Of%20the%20nearly%201.67%20million,in%20the%20past%2010%20years>
- ACT®. (2018) ACT® Tessera® technical bulletin.
- <https://www.ACT®.org/content/dam/ACT®/unsecured/documents/R1675-Tessera®-tech-bulletin-2018-08.pdf>
- Bartram, D. (2007). Increasing validity with forced-choice criterion measurement formats. *International Journal of Selection and Assessment*, 15(3), 263-272
- <https://doi.org/10.1111/j.1468-2389.2007.00386.x>
- Borkenau, P., & Ostendorf, F. (1990). Comparing exploratory and confirmatory factor analysis: A study on the 5-factor model of personality. *Personality and Individual Differences*, 11(5), 515-524. [https://doi.org/10.1016/0191-8869\(90\)90065-Y](https://doi.org/10.1016/0191-8869(90)90065-Y)
- Bratko, D., Chamorro-Premuzic, T., & Saks, Z. (2006). Personality and school performance: Incremental validity of self-and peer-ratings over intelligence. *Personality and Individual Differences*, 41(1), 131-142.
- <https://doi.org/10.1016/j.paid.2005.12.015>
- Browne M.W., & Cudeck R. (1993) Alternative Ways of Assessing Model Fit. In: Bollen K, Long J, editors. *Testing Structural Equation Models*. Sage; Newbury Park, CA: 1993. pp. 136–162.

- Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (2000). Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education. *Personality and Individual Differences, 29*(6), 1057-1068. [https://doi.org/10.1016/S0191-8869\(99\)00253-6](https://doi.org/10.1016/S0191-8869(99)00253-6)
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81-105. <http://dx.doi.org/10.1037/h0046016>
- Campion, M. C., Ployhart, R. E., & MacKenzie Jr, W. I. (2014). The state of research on situational judgment tests: A content analysis and directions for future research. *Human Performance, 27*(4), 283-310. <https://doi.org/10.1080/08959285.2014.929693>
- Cattell, H. E., & Mead, A. D. (2008). The sixteen-personality factor questionnaire (16PF). *The SAGE Handbook of Personality Theory and Assessment, 2*, 135-178. <https://doi.org/10.4135/9781849200479.n7>
- Christiansen, N. D., Burns, G. N., & Montgomery, G. E. (2005). Reconsidering forced-choice item formats for applicant personality assessment. *Human Performance, 18*(3), 267-307. https://doi.org/10.1207/s15327043hup1803_4
- Christiansen, N. D., Goffin, R. D., Johnston, N. G., & Rothstein, M. G. (1994). Correcting the 16PF for faking: Effects on criterion-related validity and individual hiring decisions. *Personnel Psychology, 47*(4), 847-860. <https://doi.org/10.1111/j.1744-6570.1994.tb01581.x>
- The College Board. (2020, September 9). *Nearly 2.2 million students in the class of 2020 took the SAT® at least once*. Newsroom.

<https://newsroom.collegeboard.org/nearly-22-million-students-class-2020-took-SAT®-least-once#:~:text=Nearly%202.2%20Million%20Students%20in,SAT®%20At%20Least%20Once%20%E2%80%93%20Newsroom>

- Conard, M. A. (2006). Aptitude is not enough: How personality and behavior predict academic performance. *Journal of Research in Personality, 40*(3), 339-346. <https://doi.org/10.1016/j.jrp.2004.10.003>
- Corstjens, J., Lievens, F., & Krumm, S. (2017). Situational judgement tests for selection. *The Wiley Blackwell Handbook of the Psychology of Recruitment, Selection and Employee Retention*, 226-246. doi:10.1002/9781118972472
- Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment, 4*(1), 5-13. <https://doi.org/10.1037/1040-3590.4.1.5>
- Cousans, F., Patterson, F., Edwards, H., Walker, K., McLachlan, J. C., & Good, D. (2017). Evaluating the complementary roles of an SJT and academic assessment for entry into clinical practice. *Advances in Health Sciences Education, 22*(2), 401-413. <https://doi.org/10.1007/s10459-017-9755-4>
- Coyle, T. R., & Pillow, D. R. (2008). SAT® and ACT® predict college GPA after removing g. *Intelligence, 36*(6), 719–729. <https://doi.org/10.1016/j.intell.2008.05.001>
- Danner, D., Aichholzer, J., & Rammstedt, B. (2015). Acquiescence in personality questionnaires: Relevance, domain specificity, and stability. *Journal of Research in Personality, 57*, 119-130. <https://doi.org/10.1016/j.jrp.2015.05.004>

- Decker, S. L., Hale, J. B., & Flanagan, D. P. (2013). Professional practice issues in the assessment of cognitive functioning for educational applications. *Psychology in the Schools, 50*(3), 300-313.
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual review of psychology, 41*(1), 417-440.
<https://doi.org/10.1146/annurev.ps.41.020190.002221>
- Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment, 18*(2), 192. <https://doi.org/10.1037/1040-3590.18.2.192>
- Duff, A., Boyle, E., Dunleavy, K., & Ferguson, J. (2004). The relationship between personality, approach to learning and academic performance. *Personality and individual differences, 36*(8), 1907-1920.
<https://doi.org/10.1016/j.paid.2003.08.020>
- FairTest. (2020, June 15). *More than half of all U.S. four-year colleges and universities will be optional for fall 2021 admission*. FairTest, The National Center for Fair and Open Testing. <https://www.fairtest.org/more-half-all-us-fouryears-colleges-and-universit>
- FairTest. (n.d.). *About FairTest*. FairTest, The National Center for Fair and Open Testing. <https://www.fairtest.org/about>
- Farsides, T., & Woodfield, R. (2003). Individual differences and undergraduate academic success: The roles of personality, intelligence, and application. *Personality and*

Individual differences, 34(7), 1225-1243. [https://doi.org/10.1016/S0191-8869\(02\)00111-3](https://doi.org/10.1016/S0191-8869(02)00111-3)

Finlayson, D. S. (1970). A follow-up study of school achievement in relation to personality. *British Journal of Educational Psychology*, 40(3), 344-348. <https://doi.org/10.1111/j.2044-8279.1970.tb02142.x>

Fiske, D. W. (1949). Consistency of the factorial structures of personality ratings from different sources. *The Journal of Abnormal and Social Psychology*, 44(3), 329. <https://doi.org/10.1037/h0057198>

Freberg, M. E., Vandiver, B. J., Watkins, M. W., & Canivez, G. L. (2008). Significant factor score variability and the validity of the WISC-III full scale IQ in predicting later academic achievement. *Applied Neuropsychology*, 15(2), 131-139. doi: [10.1080/09084280802084010](https://doi.org/10.1080/09084280802084010)

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. <https://doi.org/10.1111/j.0956-7976.2004.00687.x>

Funder, D. C. (2002). Personality psychology: Current status and some issues for the future. *Journal of Research in Personality*. [http://dx.doi.org/10.1016/S0092-6566\(02\)00515-9](http://dx.doi.org/10.1016/S0092-6566(02)00515-9)

Furnham, A., & Chamorro-Premuzic, T. (2004). Personality and intelligence as predictors of statistics examination grades. *Personality and individual differences*, 37(5), 943-955. <https://doi.org/10.1016/j.paid.2003.10.016>

- Furnham, A., Chamorro-Premuzic, T., & McDougall, F. (2003). Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. *Learning and individual Differences, 14*(1), 47-64.
<https://doi.org/10.1016/j.lindif.2003.08.002>
- Furnham, A., Monsen, J., & Ahmetoglu, G. (2009). Typical intellectual engagement, Big Five personality traits, approaches to learning and cognitive ability predictors of academic performance. *British Journal of Educational Psychology, 79*(4), 769-782. <https://doi.org/10.1348/978185409X412147>
- Galla, B. M., Shulman, E. P., Plummer, B. D., Gardner, M., Hutt, S. J., Goyer, J. P., D'Mello, S.K, Finn, A.S. & Duckworth, A. L. (2019). Why high school grades are better predictors of on-time college graduation than are admissions test scores: The roles of self-regulation and cognitive ability. *American Educational Research Journal, 56*(6), 2077-2115. [doi: 10.3102/0002831219843292](https://doi.org/10.3102/0002831219843292)
- Gatzka, T., & Hell, B. (2018). Openness and postsecondary academic performance: A meta-analysis of facet-, aspect-, and dimension-level correlations. *Journal of Educational Psychology, 110*(3), 355–377. <https://doi.org/10.1037/edu0000194>
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.)*. Boston: Allyn & Bacon.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment, 4*(1), 26–42. <https://doi.org/10.1037/1040-3590.4.1.26>
- Green, S. B. (1991). How many subjects does it take to do a regression analysis? *Multivariate behavioral research, 26*(3), 499-510.
https://doi.org/10.1207/s15327906mbr2603_7

- Gur, R. C., Richard, J., Calkins, M. E., Chiavacci, R., Hansen, J. A., Bilker, W. B., ... & Gur, R. E. (2012). Age group and sex differences in performance on a computerized neurocognitive battery in children age 8–21. *Neuropsychology, 26*(2), 251. <https://doi.org/10.1037/a0026712>
- Gur, R. C., Richard, J., Huggett, P., Calkins, M. E., Macy, L., Bilker, W. B., ... & Gur, R. E. (2010). A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: standardization and initial construct validation. *Journal of Neuroscience Methods, 187*(2), 254-262. <https://doi.org/10.1016/j.jneumeth.2009.11.017>
- Hopwood, C. J., & Donnellan, M. B. (2010). How should the internal structure of personality inventories be evaluated?. *Personality and Social Psychology Review, 14*(3), 332-346. <https://doi.org/10.1177/1088868310361240>
- Hoxby, C. M. & Turner, S. (2015). What high-achieving low-income students know about college. *American Economic Review, 105*(5), 514-517. [doi: 10.1257/aer.p20151027](https://doi.org/10.1257/aer.p20151027)
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: a Multidisciplinary Journal, 6*(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- IBM Corp. (2019). IBM SPSS Statistics for Mac, Version 26.0. Armonk, NY: IBM Corp
- Jackson, D. N., & Messick, S. (1958). Content and style in personality assessment. *Psychological Bulletin, 55*(4), 243-252. <http://dx.doi.org/10.1037/h0045996>

- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of Personality: Theory and research*, 2(1999), 102-138.
- Keiser, H. N., Sackett, P. R., Kuncel, N. R., & Brothen, T. (2016). Why women perform better in college than admission scores would predict: Exploring the roles of Conscientiousness and course-taking patterns. *Journal of Applied Psychology*, 101(4), 569. [doi: 10.1037/apl0000069](https://doi.org/10.1037/apl0000069)
- Kline, P. (1966). Extraversion, Neuroticism and academic performance among Ghanaian university students. *British Journal of Educational Psychology*, 36(1), 92-94. <https://doi.org/10.1111/j.2044-8279.1966.tb01845.x>
- Koenig, K. A., Frey, M. C., & Detterman, D. K. (2008). ACT® and general cognitive ability. *Intelligence*, 36(2), 153-160. <https://doi.org/10.1016/j.intell.2007.03.005>
- Laidra, K., Pullmann, H., & Allik, J. (2007). Personality and intelligence as predictors of academic achievement: A cross-sectional study from elementary to secondary school. *Personality and Individual Differences*, 42(3), 441-451. <https://doi.org/10.1016/j.paid.2006.08.001>
- Lawlor, S., Richman, S., & Richman, C. L. (1997). The validity of using the SAT® as a criterion for black and white students' admission to college. *College Student Journal*. 31(4), 507-515.
- Lievens, F., & Coetsier, P. (2003). Situational tests in student selection: An examination of predictive validity, adverse impact, and construct validity. *International Journal of Selection and Assessment*, 10(4), 245-257. <https://doi.org/10.1046/j.1365-2923.2002.01328.x>

- Lucido, J. A. (2018). *Understanding the test-optional movement*. In J. Buckley, L. Letukas, & B. Wildavsky (Eds.), *Measuring success: Testing, grades and the future of college admissions* (pp. 146-170). Johns Hopkins University Press.
- McCrae, R. R. (2018). Method biases in single-source personality assessments. *Psychological Assessment, 30*(9), 1160-1173. doi: [10.1037/pas0000566](https://doi.org/10.1037/pas0000566)
- McCrae, R. R. (1987). Creativity, divergent thinking, and Openness to experience. *Journal of personality and social psychology, 52*(6), 1258. doi:[10.1037/0022-3514.52.6.1258](https://doi.org/10.1037/0022-3514.52.6.1258)
- McCrae, R. R., Herbst, J. H., & Costa Jr, P. T. (2001). Effects of acquiescence on personality factor structures. *Personality and Temperament: Genetics, Evolution, and Structure, 217-231*.
- McDonald, J. D. (2008). Measuring personality constructs: The advantages and disadvantages of self-reports, informant reports and behavioural assessments. *Enquire, 1*(1), 1-19.
- McFarland, L. A., & Ryan, A. M. (2000). Variance in faking across noncognitive measures. *Journal of Applied Psychology, 85*(5), 812-821. <http://dx.doi.org/10.1037/0021-9010.85.5.812>
- Meade, A. W. (2004). Psychometric problems and issues involved with creating and using ipsative measures for selection. *Journal of Occupational and Organizational Psychology, 77*(4), 531-551. doi:[10.1348/0963179042596504](https://doi.org/10.1348/0963179042596504)

- Merz, E. L., & Roesch, S. C. (2011). A latent profile analysis of the Five Factor Model of personality: Modeling trait interactions. *Personality and Individual Differences*, 51(8), 915-919. <https://doi.org/10.1016/j.paid.2011.07.022>
- Moore, T. M., Reise, S. P., Gur, R. E., Hakonarson, H., & Gur, R. C. (2015). Psychometric properties of the Penn Computerized Neurocognitive Battery. *Neuropsychology*, 29(2), 235-246. [doi:10.1037/neu0000093](https://doi.org/10.1037/neu0000093).
- Mueller-Hanson, R., Heggstad, E. D. & Thornton, G. C. (2003). Faking and selection: Considering the use of personality from select-in and select-out perspectives. *Journal of Applied Psychology*, 88(2), 348-355. <http://dx.doi.org/10.1037/0021-9010.88.2.348>
- Mussel, P., Gatzka, T., & Hewig, J. (2016). Situational judgment tests as an alternative measure for personality assessment. *European Journal of Personality Assessment*, 34(5). <https://doi.org/10.1016/j.paid.2007.03.017>
- Muthén, L. K., & Muthén, B. O. (2007). Mplus User's Guide (8th ed.). Los Angeles, CA: Muthén & Muthén.
- Nechita, F., Alexandru, D. O., Turcu-Știolică, R., & Nechita, D. (2015). The influence of personality factors and stress on academic performance. *Current Health Sciences Journal*, 41(1), 47-61. [doi: 10.12865/CHSJ.41.01.07](https://doi.org/10.12865/CHSJ.41.01.07)
- Niessen, A. S. M., Meijer, R. R., & Tendeiro, J. N. (2017). Measuring noncognitive predictors in high-stakes contexts: The effect of self-presentation on self-report instruments used in admission to higher education. *Personality and Individual Differences*, 106, 183-189. <https://doi.org/10.1016/j.paid.2016.11.014>

- Noble, J., & Sawyer, R. (2002). Predicting different levels of academic success in college using high school GPA and ACT® composite score. *ACT® Research Report Series*. Retrieved from <https://files.eric.ed.gov/fulltext/ED469746.pdf>
- 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. <http://dx.doi.org/10.1037/0022-3514.93.1.116>
- Novikova, I. A., & Vorobyeva, A. A. (2017). Big Five Factors and academic achievement in Russian students. *Psychology in Russia*, 10(4), 95-106.
- O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual Differences*, 43(5), 971-990. <https://doi.org/10.1016/j.paid.2007.03.01>
- Olaru, G., Burrus, J., MacCann, C., Zaromb, F. M., Wilhelm, O., & Roberts, R. D. (2019). Situational Judgment Tests as a method for measuring personality: Development and validity evidence for a test of Dependability. *PloS one*, 14(2), <https://doi.org/10.1371/journal.pone.0211884>
- O'Neill, T. A., Lewis, R. J., Law, S. J., Larson, N., Hancock, S., Radan, J., ... & Carswell, J. J. (2017). Forced-choice pre-employment personality assessment: Construct validity and resistance to faking. *Personality and Individual Differences*, 115, 120-127. <https://doi.org/10.1016/j.paid.2016.03.075>
- Paulhus, D. L., & Vazire, S. (2007). The self-report method. *Handbook of Research Methods in Personality Psychology*, 1, 224-239.

- Paunonen, S. V., & Ashton, M. C. (2001). Big Five predictors of academic achievement. *Journal of Research in Personality, 35*(1), 78-90.
[doi:10.1006/jrpe.2000.2309](https://doi.org/10.1006/jrpe.2000.2309)
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*(2), 322-338.
[doi:10.1037/a0014996](https://doi.org/10.1037/a0014996)
- The Professional Standards of the National Association of School Psychologists.* (2020).
The National Association of School Psychologists.
- Rammstedt, B., & Farmer, R. F. (2013). The impact of acquiescence on the evaluation of personality structure. *Psychological Assessment, 25*(4), 1137-1145.
<http://dx.doi.org/10.1037/a0033323>
- Rammstedt, B., Kemper, C. J., & Borg, I. (2013). Correcting Big Five personality measurements for acquiescence: An 18-country cross-cultural study. *European Journal of Personality, 27*(1), 71-81. <https://doi.org/10.1002/per.1894>
- Robinson, A., (2019, December 31). *How much do SAT® Prep Courses Cost?*
PrepScholar. <https://blog.prepscholar.com/how-much-do-SAT®-prep-courses-cost>
- Sackett, P. R., Kuncel, N. R., Arneson, J. J., Cooper, S. R., & Waters, S. D. (2009). Does socioeconomic status explain the relationship between admissions tests and post-secondary academic performance? *Psychological Bulletin, 135*(1), 1. doi:
[10.1037/a0013978](https://doi.org/10.1037/a0013978)
- Sanchez-Ruiz, M. J., El Khoury, J., Saadé, G., & Salkhanian, M. (2016). Noncognitive variables and academic achievement: The role of general and academic self-

- efficacy and trait emotional intelligence. *Noncognitive Skills and Factors in Educational Attainment* , 65-85. https://doi.org/10.1007/978-94-6300-591-3_4
- Saville, P., & Willson, E. (1991). The reliability and validity of normative and ipsative approaches in the measurement of personality. *Journal of Occupational Psychology*, 64(3), 219-238. <https://doi.org/10.1111/j.2044-8325.1991.tb00556.x>
- Schinka, J. A., Kinder, B. N., & Kremer, T. (1997). Research validity scales for the NEO-PI-R: Development and initial validation. *Journal of Personality Assessment*, 68(1), 127-138. https://doi.org/10.1207/s15327752jpa6801_10
- Shewach, O. R., Shen, W., Sackett, P. R., & Kuncel, N. R. (2017). Differential prediction in the use of the SAT® and high school grades in predicting college performance: Joint effects of race and language. *Educational Measurement: Issues and Practice*, 36(3), 46-57. <https://doi.org/10.1111/emip.12150>
- Syverson, S. (2007). The role of standardized tests in college admissions: test-optional Admissions. *New Directions for Student Services*, 118, 55-70. doi: [10.1002/ss.241](https://doi.org/10.1002/ss.241)
- Trull, T. J., & Widiger, T. A. (2013). Dimensional models of personality: the five-factor model and the DSM-5. *Dialogues in clinical neuroscience*, 15(2), 135-146. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3811085/pdf/DialoguesClinNeurosci-15-135.pdf>
- Trapmann, S., Hell, B., Hirn, J. O. W., & Schuler, H. (2007). Meta-analysis of the relationship between the Big Five and academic success at university. *Zeitschrift für Psychologie/Journal of Psychology*, 215(2), 132-151. <https://doi.org/10.1027/0044-3409.215.2.132>.

- Vazire, S., & Mehl, M. R. (2008). Knowing me, knowing you: the accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology*, 95(5), 1202.
<http://dx.doi.org/10.1037/a0013314>
- Wagerman, S. A., & Funder, D. C. (2007). Acquaintance reports of personality and academic achievement: A case for Conscientiousness. *Journal of Research in Personality*, 41(1), 221-229. doi:[10.1016/j.jrp.2006.03.001s](https://doi.org/10.1016/j.jrp.2006.03.001s)
- Westrick, P.A, Marini, J.P., Young, L., NG, H., Doron, S., Shaw, E.J. (2019). *Validity of the SAT® for predicting first-year grades and retention to the second year*. The College Board®. <https://collegereadiness.collegeboard.org/pdf/national-SAT®-validity-study.pdf>
- Who Are School Psychologists? (n.d.) Retrieved from <https://www.nasponline.org/about-school-psychology/who-are-school-psychologists>
- Wilt, J., & Revelle, W. (2017). *Extraversion*. In T. A. Widiger (Ed.), *Oxford library of psychology. The Oxford handbook of the Five Factor Model* (p. 57–81). Oxford University Press.
- Ziegler, M., Danay, E., Schölmerich, F., & Bühner, M. (2010). Predicting academic success with the Big 5 rated from different points of view: Self-rated, other rated and faked. *European Journal of Personality: Published for the European Association of Personality Psychology*, 24(4), 341-355.
<https://doi.org/10.1002/per.753>

Vita

Name	<i>Allison Murray</i>
Baccalaureate Degree	<i>Bachelor of Arts, City University of New York, Queens College, Queens, NY, Major: Psychology</i>
Date Graduated	<i>May, 2015</i>
Other Degrees	<i>Master of Science, St. John's University, Queens, NY, Major: School Psychology</i>
Date Graduated	<i>May, 2019</i>