Graduate College Dissertations and Theses

Dissertations and Theses

2019

Predicting College Outcomes: A Tool for Assessing Non-Cognitive Traits in Admissions Essays

Caroline Alexandra Weaver
University of Vermont

PREDICTING COLLEGE OUTCOMES: A TOOL FOR ASSESSING NON-
COGNITIVE FACTORS IN ADMISSIONS ESSAYS

A Thesis Presented

by

Caroline Alexandra Weaver

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Master of Education
Specializing in Educational Leadership

May, 2019

Defense Date: March 19, 2019
Thesis Examination Committee:

Tammy G. Kolbe, Ed.D., Advisor
Brenda Solomon Ph.D., Chairperson
Sean M. Hurley, Ph.D.
Cynthia J. Forehand, Ph.D., Dean of the Graduate College

Abstract

Across the United States, higher education institutions increasingly employ a holistic review of prospective students' application materials. In a holistic review process, admissions offices consider a student's personal and academic context when reviewing applications for admission. A key feature of a holistic review is a student's application essay, or personal statement. However, admissions offices rarely standardize their essay review process and very little research exists regarding whether student essays predict successful outcomes in college. This paper summarizes a quality improvement study conducted within the University of Vermont Admissions Office. It examines the extent to which non-cognitive student characteristics present in student admissions essays (e.g., grit, creativity, intrinsic motivation, leadership, community engagement, cultural fluency) are correlated with pre-admission factors and subsequent college outcomes. The study involved developing a new essay scoring rubric and evaluating the usefulness of this rubric by scoring 320 undergraduate admission essays. Findings suggest that the rubric is useful in identifying evidence of non-cognitive factors in student essays, but that overall scores do not strongly correlate with pre-admissions characteristics or first-fall college GPA. The study supports the practice of holistic review and provides insight into how admissions offices can begin to operationalize the review of essays and non-cognitive factors in their admissions processes.

Keywords: admissions essay, college students, community engagement, creativity, cultural fluency, grit, holistic admissions, intrinsic motivation, leadership, love of learning, non-cognitive admissions factors, rubric validation, undergraduate admissions

Acknowledgments

I would like to thank the many individuals who supported me during the brainstorming, design, data collection, and writing stages of this project.

To Tammy Kolbe, my advisor: Thank you for your patience, guidance, and support. You helped me to turn a daunting project into a manageable step-by-step process. Thank you for always making room in your schedule to accommodate last minute questions.

To Brenda Solomon and Sean Hurley: Thank you for joining my committee mid-semester. I look forward to your valuable insights and feedback.

To Alexander Yin: You were instrumental in launching this project from its brainstorming stage to reality. Thank you for your mentorship and investment in my development as a researcher and data scientist.

To Ryan Hargraves: Thank you for your willingness to partner on this project and for the many coffee and lunch conversations that ensued. I admire your commitment to promoting equity and access within the college admissions landscape.

To Megan Nyce: Thank you for the significant investment of time you made in developing the ideal admissions essay sample. Without you, I would not have had any data to analyze. You solved the puzzle!

Table of Contents

List of Tables

List of Figures

Introduction

Many American higher education institutions employ the practice of holistic admissions (Bastedo, Howard, & Flaster, 2016). In a holistic review process, admissions offices evaluate college applications while considering a student's personal and academic context. Not all students have access to equal educational opportunities and, in an effort to improve equity among applicants and increase diversity, admissions offices have moved towards the use of context-based assessment (Clinedinst & Koranteng, 2017; Lucido, 2014). Holistic review looks at academic credentials, but also broadens the rage of factors considered by including family demographics, extracurricular activities, letters of recommendation, and student essays (Bastedo et al., 2016; Mamlet & VanDeVelde, 2011). This practice acknowledges that grade point average and standardized tests do not tell the complete student story.

With holistic admissions, there is no specific combination of factors that guarantees a student admission to an institution. Colleges and universities review many aspects of a prospective student's application and have internal processes for how they quantify, evaluate, and assign value to those factors. The National Association for College Admission Counseling (NACAC) produces an annual "State of College Admission Report" that provides up-to-date information regarding college admissions trends. In the 2017-2018 report, NACAC presented results from a survey of almost 200 colleges and universities across the United States. Admissions offices identified grades, high school curriculum, and test scores as top factors for first-time, first-year admission. Among the next most important factors was the student essay (Clinedinst & Patel, 2018).

While the relative importance schools place on the essay seems to have plateaued in the last 10 years, private colleges continue to place more importance on the essay than public institutions (Clinedinst & Patel, 2018). Additionally, institutions that are more selective in their admissions practices place more emphasis on the essay (Clinedinst & Patel, 2018).

While the essay appears to play an important role in how colleges evaluate student applications, the NACAC reports do not answer the question of how the essay is actually used in decision making. Some posit that the essay provides additional personal information that cannot be conveyed through grades and scores (Atkinson, 2001), and others argue that the essay helps to evaluate writing skills and academic fit (Kellogg & Raulerson, 2007). But the lack of research surrounding the essay's role in making admissions decisions suggests that its value may be more cultural than empirical.

Very few admissions offices have actually standardized their essay review process and no research exists on whether admissions essays are correlated with other predictors of student success in college or student outcomes once enrolled. The absence of research exposes a critical gap between the perceived and observed value of the college admissions essay. This study attempts to fill this gap by examining the extent to which non-cognitive student characteristics present in the admission essay are correlated with pre-admission factors and subsequent college academic outcomes. To do so, a new rubric for evaluating admission essays was developed, tested, and used to score 320 admissions essays from students included in the 2018 Fist-Time First-Year (FTFY) cohort at the University of Vermont.

Background

The University of Vermont (UVM) Admissions Office employs the practice of holistic admissions and requires first-year applicants to submit transcripts, standardized tests, a secondary school report, a letter of recommendation, common essay, and offers an optional supplemental essay ("First-Year Applicants," n.d.). Through a comprehensive review of these application materials, the UVM admissions office also strives to evaluate the following non-cognitive factors: 1) grit; 2) love of learning; 3) creativity; 4) leadership; 5) community engagement; and 6) cultural fluency (R. Hargraves, personal communication, January 10, 2019). While explicit in the characteristics they hope to select for, there is no evidence to suggest that any one of these characteristics should carry more weight than the other. This particular set of factors reflects UVM's local institutional priorities and as such, the admissions offices strives to admit students that display these traits.

UVM's effort to incorporate non-cognitive measures into their evaluation process reflects a larger trend in the college admissions community (Akos & Kretchmar, 2016). Non-cognitive factors are more difficult to operationalize, but are becoming increasingly important in the competitive admissions landscape. For UVM specifically, the reluctance to let quantitative factors drive the admissions process is in direct response to local educational trends. Vermont state policy is moving towards proficiency based grading in secondary schools. Just over 22 percent of students who enrolled in UVM for Fall 2018 were Vermont residents (Office of Institutional Research, n.d.) and UVM does not want to risk biasing against in-state applicants by over emphasizing quantitative standards. UVM is also looking at the future landscape of higher educations and expanding their

review process to better match the evolving applicant pool. On a national scale, UVM is one of many schools incorporating holistic review into their evaluation process, but recently the holistic process has fallen under scrutiny ("Harvard Admissions Lawsuit," n.d.). While holistic admissions attempts to increase equity in the admissions process, practices range across different schools and the process lacks transparency. A key objective for UVM's Admissions Office is to develop a new framework for operationalizing UVM's holistic admissions process by standardizing the way they review applicant essays. Designing and validating a rubric that evaluates student essays on the basis of pre-determined non-cognitive admissions factors (e.g., grit, love of learning, creativity, leadership, community engagement, and cultural fluency) is a key component of this effort.

Literature Review

Admissions Essays

In the context of the holistic admissions determination process, the application essay, or personal statement, has become a foundational component of the U.S. college application. The origin of the admissions essay can be traced back to Harvard College. The first mention of a written admissions requirement was recorded in the 1873-1874 Harvard course catalogue and was introduced in response to the perception that applicants were entering Harvard without sufficient composition skills. The 1873-74 course catalogue read, "*English Composition.* Each candidate will be required to write a short English Composition, correct in spelling, punctuation, grammar, and expression, the subject to be taken from such works of standard authors as announced from time to time"

(Brereton, 1996, p. 34). Now, most universities require an admissions essay and use writing samples not just to evaluate writing skills, but to better understand who the applicant is as a person. Students are encouraged to write personally and reveal elements of themselves that cannot be conveyed through quantitative assessment (Atkinson, 2001; Pennebaker et al., 2014). The essay also provides an opportunity for the university to assess a student's writing abilities, helping to gauge academic preparation and fit (Kellogg & Raulerson, 2007; Preiss et al., 2013). This practice is not unique to undergraduate admissions. Many graduate programs and employers require written statements from applicants, implying acceptance that a written statement can help to predict an individual's future performance (Robinson, Navea, & Ickes, 2013). The ubiquitous nature of admissions essays suggests a general consensus regarding their perceived value. However, there is a distinct absence of research-based frameworks instructing admissions committees on how to evaluate essays in the holistic decision process. Individual colleges may have specific guidelines informing holistic evaluation, but these are largely grounded in local norms and opinions as opposed to empirical evidence (Bastedo et al., 2016).

The absence of systematic research guiding essay evaluation is problematic. Essay scoring is highly susceptible to rater bias, which can have significant effects on reliability (Siu & Reiter, 2009). Additionally, few schools actually name outcomes they hope to predict by reviewing a writing sample. As such, few studies exist that explore whether admissions essays are actually helpful predictors of student academic performance post-matriculation. Those that do attempt to refine essay scoring instruments

report low inter-rater reliability and minimal correlation with student outcomes (GlenMaye & Oakes, 2002; Kretchmar, 2006).

Although very little literature exists regarding the predictive nature of admissions essays, some programs have reflected on the level of importance they assign to the essay. In a literature review of medical school admissions research, Siu and Reiter (2009) conclude that there is limited predictive value in the personal statement as a selection tool. Albanese et al. (2003) found that 41-44% of medical school applicants reported their personal statement involved input from others, with 15-51% reporting input in content development and 2-6% reporting help from professional services. These findings raise the question of what the essay is actually evaluating given that applicants receive significant outside guidance. Niessen, Meijer, and Tendeiro (2017) also explored the effect of self-presentation in applicants. When evaluating admissions processes at an undergraduate psychology program, they found that high-stakes environments (like selective admissions) inflated self-reporting of non-cognitive characteristics.

In an attempt to find meaning in the writing sample, some programs have considered the actual subject matter students write about. One baccalaureate nursing program explored the relationship between essays and attrition, finding that non-completers tended to write about nursing external to themselves, while completers described an internalization of the profession (Sadler, 2003). Non-completers were more likely to address specific characteristics of nurses they possessed, but failed to draw personal connection to the profession. Completers were more likely to discuss a personal experience with a nurse, or a family experience that exposed them to the field and generated interest in the profession. Robinson et al. (2013) tested whether differences in

the use of linguistic categories in written self-introductions for a college-level course at the start of the semester had any correlation to course performance. Results supported the possibility that relative word usage in particular linguistic categories such as punctuation and word simplicity could categorize students as "narrowed focus" or "dynamic thinkers" and predict course performance. Pennebaker et al. (2014) conducted a computerized text analysis of over 50,000 college admissions essays and found that higher college grades were associated with greater article and preposition use, referred to as "categorical language" use. However, despite these findings, the authors caution professionals from using word count analysis to aid in admissions decisions, acknowledging that enterprising students could simply game the system. Instead, they recommend the findings be used as a way to reflect on and improve the writing instruction in our educational system.

The research exploring predictive outcomes based on standardized writing exams is more substantial. The Advanced Placement (AP) exam has compared essay writing and multiple-choice questioning to evaluate a student's mastery of material and writing skills. A 1994 study found that multiple-choice AP exam scores were more highly correlated with first-year grade point average than were essay scores (Bridgeman & Lewis, 1994). The same study proposed that essay assessments produced fewer gender differences than multiple-choice tests, and that males displayed a relative advantage on multiple-choice tests. Shaw, Mattern, & Patterson (2011) reviewed components of the SAT and found that students who had relatively higher writing scores as compared to their critical reading scores on the SAT earned higher grades in their first year of college as well as in their first-year English courses. Both exams were designed to evaluate a student's

mastery of content material rather than writing skills, making the findings only partially applicable to admissions essays.

As it stands, there is no evidence in the literature suggesting that the use of a specific standardized essay rubric can help to accurately predict performance outcomes. This does not make essays void of value, but it does force colleges to question what they are actually measuring, whether the essay is the most appropriate tool for that measure, and how to effectively use the data they gather in the admissions process. It also requires colleges to consider what short and long-term outcomes they value and hope to select for.

One major concern regarding the use of essay evaluation in the holistic admissions process is the issue of bias. Standardized tests scores have been criticized for potentially displaying ethnic and cultural bias (Freedle, 2003; Santelices & Wilson, 2010). Freedle (2003) examined differential item functioning (DIF) within the SAT and found that easier verbal items on the exam favored White students over African American students. The same phenomenon was consistently present when comparing item performance between White, Hispanic, and Asian students. This relationship increased for students whose preferred language is not English. Freedle asserted that the source of this phenomenon was not simply ethnicity, but any index "that identifies a group as sharing a persistent environment that differs from the White majority English speakers" (Freedle, 2003, p. 19). Santelices & Wilson successfully replicated these findings in an updated (2010) study. Numerous studies report high correlations between standardized test scores and socio-economic status, inviting criticism that they better reflect parental financial status as opposed to student academic preparation (Crosby, Iyer, Clayton, & Downing, 2003; Zwick, 2013). The same factors that introduce bias into

standardized test scores have the potential to impact writing samples. As such, while this study explores the essay as one predictive evaluation tool for non-cognitive traits, it does not discuss the relative weight admissions offices should place on the essay. The following sections contextualize the specific non-cognitive factors UVM has focused their selection process around.

Writing Quality

Writing quality is one element that can be assessed through the review of an admissions essay. While not strictly a non-cognitive factor, the UVM admissions office acknowledges that student essays must meet a certain threshold for writing quality. In fact, assessing writing quality was the initial purpose behind inclusion of a writing sample in the college admissions process (Brereton, 1996). The essay provides the largest writing sample in the application, making it the most likely mode through which to assess writing ability (as opposed to short-answer sections throughout the application). While a great deal of educational research exists surrounding writing assessment as a tool for measuring student progress (for overview, see Huot, 1990) there is very little information available regarding the methodological ways admissions offices may define and assess writing quality. Historically, writing quality is assessed in two formats: multiple-choice tests of writing and essay tests (Breland, Bridgeman, & Fowles, 1999). In assessing writing quality, it is critical for admissions committees to establish a valid construct and to define the specific measure being isolated. The literature related to higher education tends to focus more on desired constructs rather than methods and models for assessment. Generally, the construct tested is always less comprehensive than the idealized theoretical concept (Breland et al., 1999). In a comprehensive literature review of existing writing

9

assessments used for admission in higher education, Breland et al. (1999) found that multiple-choices tests have shown to be better at assessing a student's editing ability whereas essay tests better assess the ability of an examinee to reflect on a topic and then engage in the process of conceiving, synthesizing, and presenting their own thoughts. A testing measure can also produce construct underrepresentation, which occurs when the measure does not accurately assess all important aspects of the construct. For example, a multiple choice assessment of grammar and sentence structure does not also capture critical and analytical writing ability (Breland et al., 1999). Because existing college applications do not include multiple-choice writing assessments, for the purpose of this review we will focus on the use of essay tests as an evaluator of writing quality.

The admissions essay is not a perfect analogue for an essay test. An admissions essay is not a timed assessment and as discussed previously, may involve a substantial amount of outside input, suggesting that the applicant's personal writing skills are not necessarily being reviewed. The essay is a vehicle through which a writing concept could be evaluated, but it is important to address the limitations. Research suggests that testing the comprehensive writing process is a better way to fully evaluate a writing construct (Collins & Gentner, 1980; Flower & Hayes, 1981), but assessing the writing process is not feasible in an admissions context. Studies that evaluate the writing process involve timed writing prompts, scoring submitted scratch paper, and reviewing the editing process.

Essay prompt choice is also an important factor to consider. Standardizing a prompt has the potential to increase validity, however, students vary dramatically in their experiences and exposure to topic genres (Beck & Jeffery, 2007). In a study that

explored the impact of essay prompt choice on student performance, researchers found no significant difference between performance on preferred prompts versus second choice prompts. However, students did display strong preferences and indicated more enjoyment while addressing a preferred prompt. This suggests that essay prompt choice may have more impact on perceived performance as opposed to objective performance (Powers & Fowles, 1998). Essay prompt choice is also important because it can dramatically impact the style of writing. Cumming et al. (2005) found that essay prompts based on personal experience elicited essays with significantly different characteristics than prompts that required students to integrate source texts. Personal experience prompts increased length and frequency of argument structures while the integration of source texts elicited responses with more precise vocabulary and richer academic tone. This underscores the importance of defining the desired construct and understanding the ways in which construct validity is influenced.

Construct validity can also be compromised through rating methodology. Breland & Jones (1988) found that inter-rater agreement was higher when raters scored essays in a collaborative setting rather than from remote locations. When raters could discuss approaches to interpretation in-person, there was more score convergence than if raters conducted review in isolation.

There are three main procedures for directly assessing writing quality in an essay format: primary trait, analytic, and holistic. Primary trait scoring involves the identification of one or more traits relevant to a specific writing task. The traits are related to the specific writing prompt and require a separate rubric for every unique

prompt. Primary trait scoring is generally recommended for smaller samples, as it is considered a more time consuming evaluative method (Huot, 1990).

Analytic scoring evaluates several qualities representative of good writing. These qualities are identified ahead of time and an essay is reviewed based on the presence and strength of the named qualities. Some applications of this method include Likert scale weighting of each quality to help determine a particular score. Analytical scoring was first developed by Diederich (1974) and he proposed an evaluative rubric containing qualities such as: ideas, organization, wording, flavor, usage, punctuation, and spelling. In subsequent comparison studies, analytical scoring has proven to be the most reliable of all types of scoring mechanism (Scherer, 1985; Veal & Hudson, 1983).

Holistic scoring of essays reflects a reader's general impression of the writing and does not identify multiple specific qualities. Holistic scoring is the fastest method and easiest to teach, making it a popular, albeit slightly less reliable form of writing quality assessment (Huot, 1990).

There are clearly a great number of factors influencing the methods and construct validity of writing quality assessment. For schools looking to develop their own operationalized methods, close attention must be paid to identification of desired qualities and rater training. As it stands, writing quality is typically only one component of many factors reviewed during essay evaluation. There is insufficient literature substantiating the predictive validity of written submissions, and studies that do claim predictive validity do not always include the essay rating criteria used (for example; Balogun, Karacoloff, & Farina, 1986; Berchulc, Wade, & Seidner, 1987) Grammar, syntax, and spelling as well as overall readability were two components of a more comprehensive

essay review rubric that Kirchner & Holm (1997) found predictive of in-course GPA for occupational therapy students. In another study of occupational therapy students, researchers concluded that an essay rated for "correct spelling, grammar, punctuation, clarity of statements, organization of ideas, and cohesiveness" was predictive of success in the program (Schmalz, Rahr, & Allen, 1990, p. 370). However, a great deal more evidence is needed before any significant conclusions can be drawn between specific components of writing quality in an admissions essay and college success outcomes.

Non-cognitive Factors: Grit

Grit, defined as perseverance and passion for long term goals, has recently been found to predict an array of achievement outcomes beyond traditional measures like IQ and SAT (Duckworth, Peterson, Matthews, & Kelly, 2007). The modern research surrounding grit is still in its early stages, but the concept dates back to the 1800's when psychologists were exploring the notion of human potential versus achievement. Sir Frances Galton identified the "capacity for hard labour" as a distinguishing characteristic that separated high achievers from lower achieving peers (Galton, 1869, p. 33). More recent research has centered the concept of grit in today's educational context (Duckworth et al., 2007; Duckworth & Gross, 2014). In addition, Duckworth published a New York Times bestselling book titled, "Grit: The Power and Passion of Perseverance," (2016) which popularized the concept outside of the academic realm.

Grit consists of two sub constructs: consistency of interest and perseverance of effort. Duckworth et al. (2007) identified these sub constructs through the design and validation of two self-report grit measurement scales, the twelve-item version (GRIT-O) and the eight-item short version (GRIT-S). The predictive validity of grit was assessed by

its association with higher levels of lifetime schooling among individuals of identical age (Duckworth et al., 2007). Researchers posted an online survey link and garnered participation from 1,545 participants aged 25 and older ($M = 45$ years). The sample consisted of 73% women and 27% men. An exploratory factor analysis produced a two-factor solution. The resulting 12 question Grit scale displayed high internal consistency (.85) for overall scale and for each factor (Consistency of Interest, .84; Perseverance of Effort, .78). In subsequent analysis, neither factor displayed greater predictability of outcomes, but both items considered together were consistently more predictive than either factor individually (Duckworth et al., 2007).

Grit displays similarity to self-control, but differs in that stamina is a key trait (Duckworth & Gross, 2014). Duckworth et al. (2007) assert that the "gritty" individual stays the course over an extended period of time despite temporary setbacks and disappointment. Research correlates grit with high levels of achievement in a range of fields: retention of West Point graduates through the first summer of a physically, emotionally, and mentally demanding training sequence (Duckworth et al., 2007; Duckworth & Quinn, 2009; Maddi, Matthews, Kelly, Villarreal, & White, 2012), advancement in the National Spelling Bee (Duckworth et al., 2007; Duckworth & Quinn, 2009), as well as retention and effectiveness of novice teachers (Robertson-Kraft & Duckworth, 2014).

Grit's relation to college success (defined in this context as first-year grade point average) has shown mixed results. For a sample of psychology students attending an Ivy League university, grit was associated with higher GPAs. This relationship strengthened when researchers controlled for SAT scores (Duckworth et al., 2007). In a sample of

14

Black, male students, Strayhorn (2014) found that grit added incremental validity in the prediction of college grades. Duckworth & Quinn (2009) found grit to be predictive of overall GPA one-year post-survey administration for a group of students grades seven through eleven. On the other hand, Maddi et al. (2012) failed to find a correlation between grit and first-year GPA for a sample of first-year USMA cadets, despite grit's correlation with first-year retention (Duckworth et al., 2007).

Grit's relation to college success is far from certain, but a recent study by Akos & Kretchmar (2016) served to replicate the findings of Duckworth et al. (2007), Duckworth and Quinn (2009), and Strayhorn (2014). In a sample of 209 first-year UNC-Chapel Hill students, Akos and Kretchmar (2016) found that total self-report grit scores were predictive of first-year GPA. The study builds upon the existing research suggesting that the two accepted sub constructs of grit predict GPA differently. Perseverance of effort was a superior predictor of GPA than consistency of interest. Consistency of interest was able to predict a student's likelihood to change majors, while perseverance of effort was not.

Despite grit's applicability to college success, it is important to note that grit has not been validated for high stakes conditions, nor applied as an evaluative component of college admissions essays. Social desirability bias is a pervasive problem in self-report tools and is known to increase in high-stakes environments (Schmitt et al., 2009). Even informant reports display susceptibility to enhancement biases in high-stakes environments (McDonald, 2008). Another point of increasing concern is that grit scores may be confounded by membership in certain demographic groups. Akos and Kretchmar (2016) found no significant differences in grit scores by gender, first-generation, or

15

underrepresented minority status, but observed that underrepresented minorities consistently rated themselves lower on perseverance of effort than non-minorities. As such, they state concerns about the effect internalized prejudice may have on the construct and predictive validity of grit. Others share those concerns and caution the use of grit as a single factor for admissions selection (Matteucci, Park, Patrick, Galla, & Duckworth, 2016).

Non-cognitive Factors: Intrinsic Motivation

Love of learning is another non-cognitive factor of interest to UVM, however "intrinsic motivation" is more a theoretically grounded concept and will be used as a proxy for love of learning. Intrinsic motivation displays origins in self-determination theory (SDT). SDT is an empirically-derived theory of human motivation and personality that separates motivation into contrasting categories of autonomous and controlled, or intrinsic and extrinsic (Ryan & Deci, 2000). Cognitive Evaluation Theory (CET) is a subtheory of SDT that attempts to explain variability in intrinsic motivation. The theory purports that intrinsic motivation can be cultured in educational environments that offer autonomy and positive performance feedback (Ryan & Deci, 2000). In the past 50 years, the definition of intrinsic motivation has varied slightly, but the current concept of intrinsic motivation is taken generally as a measure of liking, enjoyment, interest, curiosity, and challenge seeking (Lepper, Corpus, & Iyengar, 2005).

Higher education institutions are built upon the search for truth and pursuit of knowledge. Over time, institutions have adapted themselves to the needs of society, prioritizing more practical objectives over esoteric aims, but the historical values remain present (Vught & Westerheijden, 1994). Colleges naturally strive to attract self-motivated

16

students who value the learning opportunities college presents. However, in the competitive landscape of college admissions, identifying intrinsic motivation proves challenging.

In a study of high school science students, researchers found that individuals were more concerned with maintaining "good student identities" through their grades as opposed to developing meaningful connections to science content (Carlone, 2004). Several studies expose a troubling trend that performance serves as a prominent precursor to the development of academic interest. In other words, externally motivated factors are more likely to drive a student towards career choices than intrinsic ones (Fouad & Smith, 1996; Fouad, Smith, & Zao, 2002; Lent et al., 2003).

While both intrinsic and extrinsic factors serve to motivate learning, research supports that intrinsic motivation displays stronger correlations with academic outcomes and long-term measures of success (Trevino & DeFreitas, 2014). Intrinsic motivation may also lead to greater interest, excitement, and confidence, which in turn leads to improved performance, persistence, and creativity (Hazari, Potvin, Tai, & Almarode, 2010; Sheldon, Ryan, Rawsthorne, & Ilardi, 1997). These connections are well supported in the K-12 population. Lepper et al. (2005) administered questionnaires to students grades 3-8 measuring intrinsic and extrinsic motivation at the beginning of the school year. At the end of the year they found that intrinsic motivation positively correlated with students' grades, standardized test scores, and GPAs. Similarly, Niehaus, Rudasill, & Adelson (2012) studied Latino middle school students participating in an after school program and found that intrinsic motivation measured at the beginning of the school year positively correlated with students' GPAs at the end of the year.

17

Intrinsic motivation also displays an impact on academic factors in the college setting. Vallerand & BIssonnette (1992) have linked intrinsic motivation to higher course completion. Simons, Dewitte, & Lens (2004) found that when first-year nursing students identified a course as important to their future goals they reported higher levels of intrinsic motivation. Those with higher intrinsic motivation subsequently reported better study habits, were more excited about the course work, and persisted longer. Kaufman, Agars, & Lopez-Wagner (2008) studied intrinsic motivation at an Hispanic-serving institution and found that intrinsic motivation was positively correlated to first-quarter grades. In a regression analysis on data from 2,353 physicists and chemists, Hazari et al. (2010) found that scientists who reported a learning orientation, as opposed to a performance orientation, as their primary motivation for attending graduate school displayed more productivity in terms of total career publications and grant funding. This suggests that performance oriented motivation (extrinsic) does not sustain performance in the same way that a learning orientation (intrinsic) does. It therefore seems natural that college admissions offices would hope to select for this trait. But despite the dearth of research on intrinsic motivation, it has not been operationalized as a selection tool in the college admissions process.

The most frequently cited measurement scale for intrinsic motivation is the Echelle de Motivation en Education (EME) or it's English version, the Academic Motivation Scale (AMS). French-Canadian researcher Robert Vallerand derived the EME from SDT theory and validated it as a tool to measure intrinsic motivation (Vallerand, 1989). In 1992 the EME was translated into the English AMS and validated cross-culturally (Vallerand et al., 1992). The scale is made up of seven subscales of four

items measuring three types of intrinsic motivation (intrinsic motivation to know, to accomplish things, and to experience stimulation), three types of extrinsic motivation (external, introjected, identified regulation), and amotivation (Vallerand et al., 1992). The scale had satisfactory levels of internal consistency (.81), validating its use in educational research.

Intrinsic motivation can colloquially be described as love of learning and seems like a natural desirable quality in a college applicant. Research suggests that students who pursue their interests in an educational environment experience more productivity and excitement. But while these conclusions seem intuitive, no validated system for screening for intrinsically motivated students has been employed in the field of college admissions.

Non-cognitive Factors: Creativity

Traditional approaches to defining academic success remain linked to cognitive ability, but the power of this predictive factor cannot be evaluated without examining the appropriateness of the measured outcome (Kaufman & Agars, 2009; Schmidt & Hunter, 2004). In thinking about expanded criteria for measuring success, creativity is referenced regularly (Baird, 1977; Gose, 2005; Sedlacek, 2003). Enright & Gitomer (1989) created a tentative list of competencies thought to be critical for success in graduate studies and included creativity as a desired admissions qualification. Creativity is viewed as critical to global and economic success (Florida, 2002) and is valued in other fields of study including engineering (Cropley, 2015), business (Puccio & Cabra, 2010), and arts and sciences (Feist, 1998). Creative individuals have even reported higher levels of happiness, better physical health, and more success in entrepreneurship (Kaufman, 2006).

Creativity is sometimes viewed as an intangible strength rather than a measurable attribute (Sedlacek, 2003; Sternberg & Lubart, 1996), therefore operationalizing creativity can be difficult. Creativity relates to multiple cognitive and personal elements and there may be confluence bias with other non-cognitive variables (Sedlacek, 2003). As such, creativity has a range of operationalizations, although generally most definitions converge around the creation of a product that is viewed as useful, valued, original, and novel, or accomplishments that fit the same criteria (Dollinger, 2011; Mayer, 1999). Studies that attempt to measure creativity rely on self-ratings, or judgment of creative tasks like cartoon captioning, creative writing, or story telling (Sternberg, 2006).

Creativity is often one of the first concepts identified as missing in critiques of standardized tests, although the link between intelligence and creativity is contended (Dollinger, 2011). The view that standardized tests are unrelated to creativity is grounded in research showing little correlation between creativity and cognitive measures (Getzels & Jackson, 1962; Wallach & Wing, 1969). In a significant 45-year study of Berkeley doctoral graduates, researchers found that tested intelligence at age 27 did not predict lifetime creative accomplishments by age 72, while observer-rated intellect ratings did (Feist & Barron, 2003).

On the other hand, research exists to support the claim that standardized tests may predict creativity to a degree. In a meta-analysis, Kuncel, Hezlett, & Ones (2004) found that the MAT graduate admissions test predicted measures beyond academic performance including performance in jobs and creativity. A 1992 study of Berkeley doctoral students also found that MAT scores correlated with professor ratings of "creative quality of students" (Gough, 1992). A more recent study followed young scholars who scored in the

top .01% on the SAT. More than 20 years later, these individuals displayed highly creative individual and occupational accomplishments (Lubinski, Benbow, Webb, & Bleske-Rechek, 2006). The research does not support the more extreme claim that standardized tests fully bias against creativity, but studies that have found correlations are limited by their focus on self-ratings or judge-ratings of creativity (Dollinger, 2011).

With mixed schools of thought, why expand admission criterion to include creative achievements? Empirical studies have exposed highly exciting findings when restructuring selection criteria to include creative measures. After designing a predictive test that focused on non-cognitive factors like creativity, Sternberg (2006, 2009, 2010) found he could dramatically increase prediction of first-year university performance while simultaneously reducing ethnic group differences. Arguably one of the most vocal critics of the traditional admissions process, Sternberg (2004, 2010) has applied his theory of successful intelligence to advocate for the use of essential but unmeasured constructs like creativity, practical intelligence, and wisdom in the admissions realm. Sternberg asserts that in all walks of life people need "(a) creativity to generate new and exciting ideas, (b) analytical intelligence to evaluate whether their (and others') ideas are good ideas, and (c) practical intelligence to execute their ideas and to persuade others of their value" (Sternberg, 2009, p. 1). Sternberg acknowledges the reasonable predictive validity of the SAT for projecting undergraduate performance, but argues for the value of augmenting this assessment with measures that evaluate a range of creative and practical skills (Sternberg, 2010).

Sternberg & the Rainbow Project Collaborators (2006) developed their own set of supplemental admissions measures (The Rainbow Measures) where students respond to

21

hypothetical situations and produce creative work (Pretz & Kaufman, 2017). Creative assignments, such as captioning New Yorker cartoons and dictating short creative stories were rated by judges on a Likert-scale rubric for novelty, quality, and task appropriateness (Sternberg, 2006, 2009, 2010). Results displayed a strong correlation (.77) between latent creativity factors and SAT-Verbal scores. They also found that The Rainbow Measures combined were twice as powerful in predicting undergraduate GPA than the SAT and showed far fewer ethnic differences in scores across categories (Sternberg, 2010). This was considered a successful research project with strong implications for increasing equity in the undergraduate admissions process.

Sternberg ultimately went on to implement the procedures at Tufts University as part of The Kaleidoscope project. For two years, Tufts applicants were provided optional supplemental essay prompts that were more creative than traditional prompts. Sternberg (2009) describes the optional creative questions:

> As examples of items, a creative question asked students to write stories with titles such as "The End of MTV" or "Confessions of a Middle-School Bully." Another creative question asked students what the world would be like if some historical event had come out differently, for example, if the Nazis had won World War II. Yet another creative question, a nonverbal one, gave students an opportunity to design a new product or an advertisement for a new product. (p. 283)

Essays were scored in the same fashion as The Rainbow project and no meaningful differences across ethnic groups were found. Tufts received anecdotal feedback from applicants expressing appreciation for the opportunity to display more creativity. Tufts

sent messages to students, parents, and secondary schools emphasizing their new focus

and that year applications from African-Americans and Hispanic-Americans increased

significantly. That same year, African-American admissions increased 30% and

Hispanic-American admissions increased by 15% (Sternberg, 2009). The implications for

increasing equity and access in the admissions realm are exciting.

Building on Sternberg's work, Pretz & Kaufman (2017) challenged the validity of

current predictive measures of success in admissions by comparing selection criteria to

creative task performance for the same group of students. Results showed that traditional

admissions criteria were only weakly related to creativity suggesting that current

processes do not reward creative students. Dollinger (2011) has argued that creative

students are not necessarily penalized by the process, but when considering how to

broaden the definition of "success" schools need to consider what type of applicants they

are selecting for and what measures they want to place value on. While the research on

creativity as an admissions criteria is exciting, the link between creativity and

demonstrated student outcomes is still weak.

Non-cognitive Factors: Leadership

Leadership is regularly used as an evaluative component in college admissions

and therefore differs from the other non-cognitive factors discussed in this review.

Leadership can be described as performing a managing role in a group, motivating others,

or coordinating groups and tasks (Ryan, 2017). While student leadership may involve

formal leadership titles in school or community extracurricular activities (president, vice

president, chair, vice-chair, captain, co-captain, founder, etc.) leadership can also be more

informal. Traditional definitions of leadership often depict leaders as competing to rise

within a hierarchy, while newer theoretical perspectives define leadership as more collaborative and inclusive (Sax & Harper, 2007).

Although leadership experience is often listed as a desirable trait in college admissions, as is becoming increasingly thematic, few studies examine leadership experience as a predictive variable for college success. In a study of academically at-risk students, Mattson (2007) explored pre-college variables that correlate with first-year college outcomes. High school GPA, gender, and leadership experience emerged as the strongest predictors of first-semester GPA and fist-year GPA. Leadership was specifically defined as peer related leadership, and only individuals with formal leadership positions (president, vice-president, chair, etc.) were classified as leaders (Mattson, 2007). Another 2007 study concluded that high school leadership activities predict higher college attendance rates for all demographic groups. The study went on to look specifically at Hispanic students and found that high school leadership activities predict a higher probability of obtaining a college degree for Hispanic students whose first language is not English (Lozano, 2008). The relationship becomes even stronger for students whose first post-secondary experience is at a 2-year college. In a study of White, male students, Kuhn & Weinberger (2005) found that high school leaders were more likely to occupy managerial positions as adults and ultimately commanded higher wages. Various theories exist to help explain the relationship between leadership and college outcomes, but no reasoning is conclusive. There tends to be overlap between high school extracurricular involvement research and high school leadership research. Barron, Ewing, & Waddell (2000) argue that sport participation (regardless of leadership on the team)

teaches discipline, confidence, and teamwork, which helps to explain the correlations between participation and increased educational attainment.

The Non-Cognitive Questionnaire (NCQ) is an admissions tool that was developed in 1984 that assess college preparedness based on non-cognitive factors. The NCQ purported to decrease admissions bias against Black students by deemphasizing standardized test scores (Tracey & Sedlacek, 1984). Authors reported correlations between the eight dimensions of the NCQ and first-year college grades (Successful Leadership Experience is one of the eight dimensions of the NCQ). But while the NCQ presents an appealing tool and is cited widely, it is also heavily critiqued. In 2007, Thomas, Kuncel, & Credé conducted a meta-analytic review of the validity scores of the NCQ, finding none of the scales to be predictors of GPA or persistence in college (Thomas, Kuncel, & Credé, 2007)

The literature surrounding college-level leadership involvement is slightly more robust, albeit equally inconclusive. A 2009 study exploring both cognitive and non-cognitive predictors of college success concluded that while cognitive factors like high school GPA and SAT/ACT scores were the strongest predictors of cumulative college grade point average, strong correlations between college leadership opportunities and graduation were also observed (Schmitt et al., 2009). Cress et al. (2001) critique the fact that while most college mission statements claim to develop leadership skills, minimal attention is paid to formal leadership curricula. A longitudinal study of 875 students across 10 institutions revealed that leadership participation in college had a positive impact on developmental and growth outcomes. Participants in leadership activities were

more likely to report growth in their commitment to civic responsibility, conflict resolution skills, as well as multi-cultural awareness (Cress et al., 2001).

Student leadership research is not devoid of equity concerns. While Lozano (2008) presented positive correlations between high school leadership experience and college outcomes for Hispanic students, he also highlighted the fact that Hispanic students were less likely than White students to hold leadership positions. However, once unfavorable demographic characteristics were controlled for (lower SES, test scores, attending schools with fewer leadership opportunities, etc.), Hispanic students participating in extracurricular activities were slightly more likely than White students participating in the same activities to hold leadership positions (Lozano, 2008). Hoffman (2002) also cautions against overemphasizing the importance of college leadership based on the fact that most positions of value reflect the dominant culture on campus. Students who achieve leadership positions are likely those whose culture aligns with the dominant culture (Hoffman, 2002).

Multiple positive outcomes are reported to exist for students participating in high school and college leadership activities. However, there is a great deal left to learn about exactly how leadership translates to academic or professional success and whether leadership can be operationalized as a measure for college admissions. The research also raises concerns about unequal access to leadership and generally cautions practitioners from overemphasizing leadership participation as an isolated factor for selection, lest it continue to accentuate some of the biases that already exist within cognitive admissions factors.

Non-cognitive Factors: Community Engagement

Community engagement as a non-cognitive factor encompasses a wide range of definitions. Community engagement can be thought of as building and maintaining relationships with individuals in the school and broader community. In a college application, community engagement can present through involvement in volunteer or extracurricular activities (clubs, organizations, sports teams, etc.). The broad definition ultimately draws upon multiple fields of research on extracurricular engagement and volunteerism.

Colleges have traditionally emphasized, and continue to reinforce a mission of helping students develop both socially and ethically (Sullivan & Rosin, 2008). The number of American students participating in volunteer activities has steadily increased over time (Hart, Matsuba, & Atkins, 2008). The perception amongst students applying to college is that volunteerism is viewed favorably, which has contributed to the idea of "resume building activities." Research has correlated college service participation with academic development and civic engagement (Astin & Sax, 1998), yet the research surrounding mandatory versus voluntary participation is mixed. Stukas, Snyder, & Clary (1999) found that college students who were required to participate in "mandatory volunteerism" programs reported reduced intentions of future volunteering. Effects were less strong for students who felt positively about volunteerism before entering the mandatory program. In a contradicting study, researchers followed students at a school before and after the implementation of a mandatory volunteerism program. Both groups of students experienced the same gains in civic interest regardless of whether participation was mandatory or voluntary (Metz & Youniss, 2003).

Studies that do explore positive outcomes associated with volunteerism tend emphasize psychological wellness as opposed to academic outcomes and focus more on older populations. In a longitudinal study of older adults, Piliavin & Siegl (2007) found that consistency of volunteering and diversity of participation significantly correlated to well-being and self-reported health. Similarly, Morrow-Howell, Hinterlong, Rozario, & Tang (2003) found that older adults who volunteer report higher levels of well-being. Some studies even suggest that the benefits of volunteering are weaker or nonexistent for adults in their midlife (Van Willigen, 2000).

Although volunteerism in younger populations is studied less frequently, positive outcomes associated with youth volunteerism have been observed. Hart, Donnelly, Youniss, & Atkins (2007) found that community service in high school was a strong predictor of adult voting and volunteering. Other studies corroborate the link between young volunteerism and continued volunteerism (Sax, Astin, & Avalos, 1999; Wilson & Musick, 1997). Engagement in volunteering during young adulthood has also been correlated with feelings of personal efficacy (Reeb, Katsuyama, Sammon, & Yoder, 1998). One longitudinal study of college students found positive effects between volunteerism and personal growth, purpose, and life satisfaction (Bowman, Brandenberger, Lapsley, Hill, & Quaranto, 2010).

Community engagement can also occur through extracurricular participation. There is a great deal of research exploring the associations between extracurricular participation and academic outcomes. Many colleges explicitly encourage students to become actively engaged in their school and local communities through extracurricular involvement (Dumais, 2008). However, the way in which extracurricular participation

relates specifically to college success (i.e. college GPA, graduation rates) is less understood. Like all factors considered in college admissions, individual schools may develop different reasons for ascribing value to extracurricular involvement, but past research does indicate that school-sponsored extracurricular activities are positively associated with academic outcomes (Eccles & Barber, 1999; Gerber, 1996). For students participating in extracurriculars, researchers have documented higher standardized test scores (Dumais, 2008; Fredricks, 2012) and higher high school grades (Fredricks, 2012; Gardner, Roth, & Brooks-Gunn, 2008). In a longitudinal study of high school students, Lleras (2008) found that students who participated in extracurricular activities in high school experienced higher educational attainment and earnings.

Research suggests that not all extracurricular activities are created equal and that participation may reach a point of diminishing returns (Fredricks, 2012). One theory is that the time allocated towards intense extracurricular involvement may detract from time spent on academics (Cotter, Pretz, & Kaufman, 2016). Furthermore, different extracurricular domains (i.e. athletics, academics, arts, etc.) do not uniformly relate to academic achievement (Cotter et al., 2016). In a study that examined participation (not level of intensity), Hunt (2005) found that high school students with higher grades were more likely to participate in academic and artistic extracurricular endeavors whereas athletic and vocational clubs were more likely to attract students with lower grades. Eitle (2005) found that participation in individual sports and specific team sports correlated with higher test score achievement for males and females. Playing baseball/softball, football, and basketball was negatively associated with male tests scores, but showed no association with female test scores. Eitle (2005) suggests that some of these differences

can be attributed to selection bias, but that achievement benefits for sports seem more consistent for females.

It is important to discuss the class achievement gaps associated with extracurricular engagement. The literature surrounding extracurricular engagement and access suggests that the measure of extracurricular participation may serve as a proxy for socioeconomic class (Snellman, Silva, Frederick, & Putnam, 2015). Extracurricular activities often meet outside of school hours, require transportation, and are increasingly associated with participation fees, especially non-school sponsored activities. Students with sibling caretaker responsibilities or from less resourced families may not have equal access to extracurricular opportunities and the literature supports these concerns (Snellman et al., 2015). Children from upper-middle-class families are much more likely to join school clubs and participate on sports teams than their working-class peers (Marsh & Kleitman, 2002; Marsh, 1991; Snellman et al., 2015). Lareau (2011) found that elementary school-aged children from middle-class backgrounds were more likely to be enrolled in structured extracurricular activities than their working-class peers, resulting in educational benefits. The working-class children were more likely to spend time with friends or watch television after school and had fewer adult interactions. Several studies have found a negative association between student employment and academic outcomes, with negative associations increasing for students working more than 20 hours per week (Marsh, 1991; Safron, Schulenberg, & Bachman, 2001). In a study of South Texas high school students, Weller, Kelder, Cooper, Basen-Engquist, & Tortolero (2003) found that tobacco, alcohol, and marijuana use increased as hours of work increased. Although extracurricular engagement may correlate with multiple positive academic outcomes,

colleges need to consider the populations they are excluding before over ascribing value to this particular measure.

Community engagement as a non-cognitive factor has a broad reaching definition. While there are numerous studies linking volunteerism and extracurricular participation to positive personal outcomes, no studies were found that described validated processes through which colleges select for this trait in the essay or otherwise. Nor is there a perfect understanding of how community engagement directly links to college academic outcomes. Like most non-cognitive factors, the value is more implied than substantiated.

Non-cognitive Factors: Cultural Fluency

The shift to holistic admissions represents a desire to increase diversity on college campuses (Lucido, 2014). Holistic college admissions is based on the premise that a diverse student body increases the number of interactions between individuals from different backgrounds and enriches the learning environment (Pike, Kuh, & Gonyea, 2007). While holistic admissions is not exactly synonymous with affirmative action, in recent years more schools have adopted affirmative action policies to help increase diversity on campus (Palmer, 2001). There is an overwhelming body of research indicating connections between study body diversity and greater openness to and understanding of diverse people (Chang, 2001; Henderson-King & Kaleta, 2000; Pascarella, Palmer, Moye, & Pierson, 2001; Pike et al., 2007; Whitt, Edison, Pascarella, Terenzini, & Nora, 2001).

The University of Vermont's definition of cultural fluency describes a student that "actively seeks engagement with individuals who hold different identities and who values a diversity of experiences and opinions" (R. Hargraves, personal communication, January

31

10, 2019). The concept of cultural fluency reflects a desire to attract and admit students who are open to diverse experiences. Whitt et al. (2001) found that students who engaged in challenging conversations and encountered a more diverse population of students and student experiences in college displayed more growth in their openness to diversity and challenge. The most significant (albeit not surprising) positive effect was in students who reported the most openness to diversity and challenge *before* entering college. So while all students serve to benefit from diverse campuses, students who seek diverse experiences in high school display the most growth in college.

With diversity being such a focus on college campuses, it makes sense that institutions would strive to select for students that display openness to diversity. This concept has also gained footing in graduate school admissions. The Association of American Medical Colleges (AAMC) has articulated the desire to include non-cognitive factors ("personal competencies") in medical school admissions. The AAMC underwent a multi-year process to identify core personal competencies of value. One of the nine core competencies was "cultural competency" defined as "Demonstrates knowledge of social and cultural factors that affect interactions and behaviors; shows an appreciation and respect for multiple dimensions of diversity; recognizes and acts on the obligation to inform one's own judgment; engages diverse and competing perspectives as a resource for learning, citizenship, and work; recognizes and appropriately addresses bias in self and others; interacts effectively with people from diverse backgrounds" (Koenig et al., 2013, p. 607). Based on a survey of 98 admissions officers at MD-granting schools, cultural competency was listed as a 3.7 (in between important and very important) on an importance scale ranging up to 5 (extremely important). Although the AAMC has

identified cultural competency and other personal competencies as important for admission, they cite the need for more research to develop a validated selection tool. This reflects a theme in the literature; while virtually all colleges acknowledge the importance of diverse environments and open minded applicants, no schools have operationalized the concept.

Study Purpose and Scope

UVM's desire to incorporate the evaluation of non-cognitive factors into their admissions process reflects a growing trend in the field of admissions. Previous research has established many positive relationships between non-cognitive factors and student outcomes. The research even suggests that prioritizing non-cognitive traits in the college admissions process could help in reducing the cultural and ethnic biases perpetuated by standardized test scores. But while many of the non-cognitive factors can be widely accepted as desirable traits to select for in an applicant, there are very few empirical studies that operationalize the factors in the context of admissions. The relationship between non-cognitive factors, student admissions essays, and subsequent college outcomes has yet to be determined with any decisiveness.

The goal of this study is to create and validate an essay evaluation rubric that operationalizes the following non-cognitive traits: 1) grit; 2) intrinsic motivation; 3) creativity; 4) leadership; 5) community engagement; and 6) cultural fluency. This study describes the development and testing of a tool that can be used by professionals in scoring admissions essays and attempts to answer the questions:

1. To what extent can non-cognitive factors be operationalized to assist in the review of undergraduate application essays?

2. To what extent can non-cognitive factors displayed in undergraduate application essays predict pre-admissions factors (high school grade point average, ACT/SAT scores) and college outcomes (first-semester college grade point average)?

Data & Methods

Study Overview

This study involved the development and validation of an essay evaluation tool designed to operationalize six non-cognitive factors (e.g., grit, intrinsic motivation, creativity, leadership, community engagement, and cultural fluency). The tool was revised through field interviews and subsequently tested on a sample of 320 admissions essays written by first-year students at the University of Vermont. The sample was representative across a wide range of key student characteristics including application type (early action vs. regular action), gender (female vs. male), race/ethnicity (White vs. student of color) first-generation college status (first-generation vs. not first-generation), high school GPA (high HS GPA vs. low HS GPA), and standardized test scores (high ACT/SAT vs. low ACT/SAT).

Rubric Development

Rubric development was a two-phase process. Phase 1 included literature review and preliminary design while Phase 2 involved field testing and revision.

The UVM Admissions Office provided the list of non-cognitive factors they hope to select for in their admissions process: grit, intrinsic motivation, creativity, leadership, community engagement, and cultural fluency. A literature review was conducted for each of the non-cognitive factors and existing definitions, constructs, and subconstructs were recorded. The preliminary rubric was designed containing four levels of performance for each non-cognitive factor (no evidence, low 1-2, middle 3-4, high 5-6). Definitions of the non-cognitive factors were obtained from studies that attempted to operationalize the same factors. Written descriptions for each level of performance were either modeled off examples in the literature or generated from the definitions. A complete copy of the preliminary rubric can be found in Appendix A. The approach to Phase 1 development was to intentionally include more information than could be retained in a functional rubric.

The preliminary rubric was field tested during Phase 2. Five experts in the field of college admissions were recruited from UVM as well as peer and peer-aspirant institutions. Each expert participated in an hour-long semi-structured interview to provide feedback on the preliminary rubric[1]. A copy of the interview protocol can be found in Appendix B. During the interview, each expert was asked to define the six non-cognitive traits in their own words and provide performance definitions. At the end of the interview, participants were shown the preliminary rubric and asked to provide direct feedback on the structure and definitions. Once all five interviews were completed, a systematic review of interview notes was conducted and emerging themes were recorded.

---

[1] This phase of the project was considered quality improvement and deemed "not research" by the University of Vermont's Institutional Review Board.

All of the interviewees agreed that the rubric needed to be succinct, simple, and user-friendly. Economy and efficiency should be prioritized if the rubric was going to be integrated into the application review process; a clunky rubric that slowed down review would not be tolerated. Interviewees indicated that the number of subconstructs for each non-cognitive factor should be reduced down to the ideas most representative of the overall definition, or eliminated entirely. Additionally, all interviewees agreed that a writing quality threshold must be met for them to recommend a student for admission. A student may display strengths in many non-cognitive areas, but it was critical that they simultaneously displayed competency in their writing abilities. Initially, a grammar/style category was not included in the rubric, but it was added in response to the interviewee feedback. All of the non-cognitive traits were reduced down to one comprehensive definition and subconstructs were either removed or integrated into the definition and performance standards. A finalized version of the rubric is presented in Figure 1.

Figure 1

Final Essay Evaluation Rubric

	No Evidence (0)	Low (1, 2)	Mid (3, 4)	High (5, 6)
Grammar and Style • Displays appropriate word use, sentence structure, grammar, spelling, and punctuation. • Writes coherently and in an organized manner	Essay contains innumerable grammar, spelling, and punctuation errors. Essay organization is illogical.	Essay contains frequent errors in grammar, spelling, and punctuation. Essay organization is ineffective in building an argument.	Essay is mostly free of grammatical, spelling, and punctuation errors. Essay is adequately organized in a way that builds an argument.	Essay is free of grammatical, spelling, and punctuation errors. Student effectively organizes ideas and builds a logical coherent argument.
Grit • Perseverance of effort despite setbacks and challenges • Articulates the ways in which challenges and failure have provided value	Student has not experienced adversity or hardship or is unable to persevere when challenges arise.	Student has experienced some hardship or setbacks. Student displays little ability to persevere amidst adversity or does not attempt to draw conclusions about the value of the experience(s).	Student has experienced some hardship or setbacks and has attempted to persevere amidst adversity. Student attempts to articulate how the experience(s) have shaped them.	Student has experienced significant hardship or setbacks and consistently displays perseverance amidst adversity. Student displays self-awareness in their ability to articulate how the experience(s) have shaped them.
Intrinsic Motivation • Pursues an activity for itself and the pleasure and satisfaction derived from participation • Pursues an academic passion beyond what is expected or prescribed • Values improvement over external awards and praise	Student does not display an interest in academic or extracurricular pursuits beyond what is expected of them by family and teachers.	Student displays interest in academic and/or extracurricular pursuits, but does not sustain the interest and appears to be more motivated by outside recognition and awards than genuine curiosity.	Student displays a genuine interest in academic and/or extracurricular pursuits, but may rely on family or teachers to guide their learning and extension.	Student displays genuine interest in academic and/or extracurricular pursuits and employs self-guided learning tactics to extend their interest beyond what is expected of them by family and teachers. Student's motivation is internal and does not rely on outside recognition or awards.

Final Essay Evaluation Rubric, Continued.

	No Evidence (0)	Low (1, 2)	Mid (3, 4)	High (5, 6)
Creativity • Approaches problems and ideas in new ways • Escapes the bounds of conventional thinking • Distinguishes themselves from others through creative writing or creative extracurricular pursuits	Student does not display unique problem solving approaches. Student either fails to distinguish themselves from conventionality in their pursuits or takes risks in their writing that display poor judgment.	Student attempts to approach a problem from a new perspective. Student's writing style or activities diverge from conventionality, but fails to distinguish themselves from others.	Student attempts to approach a problem from a new perspective. Student's writing style or activities diverge from conventionality in a positive way.	Student displays a novel problem solving approach and brings a fresh or unique perspective to an idea/issue. They diverge from conventionality through their writing style or activities in a way that is appealing and unique.
Leadership • Demonstrates the skills to motivate others • Serves as the main representative for a group either officially or unofficially (may include leadership within family structure) • Actions provide guidance to others	Student displays minimal to no involvement in group activities and does not hold leadership positions	Student displays limited involvement in group activities. Evidence of leadership is either limited to title or presents as superficial.	Student displays involvement in group activities. Evidence of leadership is demonstrated by some ability to motivate others. Student may hold multiple leadership titles or responsibilities.	Student displays committed involvement to a number of group activities and is skilled in motivating others. Student may hold multiple leadership roles, but commitments emphasize depth over breadth.
Community Engagement • Builds and maintains relationships with individuals in the school and broader community • Demonstrates commitment to improving the broader community	Student does not display a commitment to improving their school or broader community. Any participation in community activities appears self-serving.	Student displays limited commitment to improving their school or broader community. Student has built some relationships, but does not actively foster them.	Student displays some commitment to improving their school or broader community. Student displays the ability build relationships. Participation does not appear self-serving	Student displays substantial commitment to improving their school or broader community. Strong evidence of relationship building and an appreciation for service exists. Participation does not appear self-serving.

	No Evidence (0)	Low (1, 2)	Mid (3, 4)	High (5, 6)
Cultural Fluency • Actively seeks engagement with individuals who hold different identities • Values diversity of experiences and opinions	Student does not discuss an understanding of identify and shares no evidence of engaging with individuals who hold different identities.	Student displays an unsophisticated understanding of identity and rarely engages with others who hold different identities.	Student displays an understanding of identity and has occasionally engages with others who hold different identities.	Student displays a complex understanding of identify and actively seeks to engage with others who hold different identities.

Essay Selection

The goal when designing the essay sample was to obtain broad representation across a number of key applicant characteristics including application type (e.g., early action vs. regular action), gender, race/ethnicity, and first-generation college status. Additionally, to test the rubric's effectiveness, it was important to have a broad distribution of high school GPA and ACT/SAT scores within the sample.

The essay selection process occurred in two stages. First, the list of students to be included in the sample was identified, and second, the essays corresponding with each student in the sample were obtained.

The UVM Office of Institutional Research (OIR) identified the students meeting selection criteria. Figure 2 provides a schematic of the steps involved in selecting the sample. The initial population included the Fall 2018 Fist-Time First-Year (FTFY) cohort at the University of Vermont. Virtually all of the students in the Fall 2018 FTFY cohort took the SAT after March of 2016 and therefore had scores reflecting the new assessment scale (total scores out of 1600, not 2400). Only students who took the exam with the new

scoring system were retained. Reported SAT scores were converted to a comparable ACT score to create a universal standardized test score. If a student reported both SAT and ACT scores, the higher of the two scores was used. Twenty students were removed from the population dataset because they either did not have high school GPA data, or they took the older version of the SAT and did not take the ACT. After these adjustments were made, the starting population contained 2,511 students.

The population was then divided into eight groups (Step 1, Figure 2): 1) Student of Color, Early Action, Female; 2) Student of Color, Early Action, Male; 3) Student of Color, Regular Action, Female; 4) Student of Color, Regular Action, Male; 5) White, Early Action, Female; 6) White, Early Action, Male; 7) White, Regular Action, Female; and 8) White, Regular Action, Male.

Next, quartiles were defined for converted ACT scores (Table 1) and high school GPA (Table 2). A 4x4, 16-category matrix was created to display the intersection of these quartiles (Table 3). To ensure the sample included representation across academic achievement, 10 students within each of the 16 matrix categories were selected from the eight groups (Step 2, Figure 2). The result was smaller, more balanced population ($n =$ 1,280) from which to create the sample.

Finally, 20 students were selected in each of the 16 matrix categories from the more balanced population (Step 3, Figure 2). The result was a final sample of 320 students. Table 4 displays the distribution of the original population and final sample for the 16 matrix categories.

Table 5 provides a summary of population and sample characteristics. The final sample had 20 essays within each of the 16 GPA/ACT quartiles and was roughly

balanced by application type (early action = 52.2%, regular action = 47.8%) and gender (female = 53.1%, male = 46.9%). The original population was much more skewed (67.6% early action, 62.7% female). The final sample could not be evenly balanced by race/ethnicity, but the final composition increased the representation of students of color from 11.7% to 30%. Additionally, 15.9% of the students in the sample were first-generation college students compared to 11.9% in the original population.

The UVM Office of Institutional Research (OIR) assigned a randomly-generated identification number (ID) to each of the 320 students in the sample. A Microsoft Excel file containing both the UVM student ID and randomly generated ID was securely transferred to the UVM Admissions Office. Using this list, the UVM Admissions Office downloaded and printed a copy of each corresponding admissions essay from their database. Admissions staff labeled each essay with the randomly generated ID number, redacted any personally identifying information, and emailed scanned copies of the essays to OIR. OIR subsequently created a new Excel file, adding relevant student information (e.g., applicant characteristics, high school GPA, ACT score, first-fall UVM GPA), and deleted the student's UVM ID number. The de-identified essays and de-identified file of student data was then sent to the researcher for coding.[2,3]

---

[2] Approval for this stage of the project was obtained through the University of Vermont's Institutional Review Board.

[3] UVM's Office of Institutional Research maintained a crosswalk of the assigned randomly-generated identification numbers and UVM student identification numbers. The crosswalk file will be deleted by OIR at the study's conclusion.

Figure 2
Designing the Essay Sample

Step 1

Divide the initial population into 8 groups by student characteristic

Step 2

For every group, select 10 students from each of the 16 ACT/GPA quartile bins

Step 3

From the new population, Select 20 students from each of the 16 ACT/GPA quartile bins

Initial Population
2018 first-time, first-year UVM cohort
$n = 2,511$

| SOC, EA, F | SOC, EA, M | SOC, RA, F | SOC, RA, M | W, EA, F | W, EA, M | W, RA, F | W, RA, M |

Quartiles	ACT 1	ACT 2	ACT 3	ACT 4
GPA 1	10	10	10	10
GPA 2	10	10	10	10
GPA 3	10	10	10	10
GPA 4	10	10	10	10

smaller, more balanced population
$n = 1,280$

Quartiles	ACT 1	ACT 2	ACT 3	ACT 4
GPA 1	20	20	20	20
GPA 2	20	20	20	20
GPA 3	20	20	20	20
GPA 4	20	20	20	20

Final Sample
$n = 320$

Note. SOC = Student of Color, W = White, EA = Early Action, RA = Regular Action, M = Male, F = Female

Table 1
Average ACT score by ACT Quartile Grouping

Quartile	Average ACT Score	n
1	22.44	579
2	25.52	570
3	28.00	790
4	31.45	572

Table 2
Average High School GPA by High School GPA Quartile Grouping

Quartile	Average High School GPA	n
1	3.13	629
2	3.54	628
3	3.79	626
4	3.98	628

Table 3
Design of 16-category Matrix

Quartiles	ACT 1	ACT 2	ACT 3	ACT 4
GPA 1	GPA1, ACT1	GPA1, ACT2	GPA1, ACT3	GPA1, ACT4
GPA 2	GPA2, ACT1	GPA2, ACT2	GPA2, ACT3	GPA2, ACT4
GPA 3	GPA3, ACT1	GPA3, ACT2	GPA3, ACT3	GPA3, ACT4
GPA 4	GPA4, ACT1	GPA4, ACT2	GPA4, ACT3	GPA4, ACT4

Table 4

Distribution of Initial Population versus Sample by Quartile Grouping

Final Quartile Category	Population		Sample	
	n	%	*n*	%
GPA1, ACT1	203	8.10%	20	6.30%
GPA1, ACT2	155	6.20%	20	6.30%
GPA1, ACT3	174	6.90%	20	6.30%
GPA1, ACT4	97	3.90%	20	6.30%
GPA2, ACT1	165	6.60%	20	6.30%
GPA2, ACT2	173	6.90%	20	6.30%
GPA2, ACT3	194	7.70%	20	6.30%
GPA2, ACT4	96	3.80%	20	6.30%
GPA3, ACT1	130	5.20%	20	6.30%
GPA3, ACT2	128	5.10%	20	6.30%
GPA3, ACT3	217	8.60%	20	6.30%
GPA3, ACT4	151	6.00%	20	6.30%
GPA4, ACT1	81	3.20%	20	6.30%
GPA4, ACT2	114	4.50%	20	6.30%
GPA4, ACT3	205	8.20%	20	6.30%
GPA4, ACT4	228	9.10%	20	6.30%

Table 5

Population and Sample Distribution by Applicant Characteristic

Characteristic	Population		Sample	
	n	*%*	*n*	*%*
SOC, EA, Female	113	4.5%	39	12.2%
SOC, EA, Male	43	1.7%	12	3.8%
SOC, RA, Female	96	3.8%	31	9.7%
SOC, RA, Male	42	1.7%	14	4.4%
White, EA, Female	987	39.3%	53	16.6%
White, EA, Male	554	22.1%	63	19.7%
White, RA, Female	378	15.1%	47	14.7%
White, RA, Male	298	11.9%	61	19.1%
Application Type				
EA	1,697	67.6%	167	52.2%
RA	814	32.4%	153	47.8%
Gender				
Female	1,574	62.7%	170	53.1%
Male	937	37.3%	150	46.9%
Race/Ethnicity				
White	2,217	88.3%	224	70.0%
SOC	294	11.7%	96	30.0%
First-Gen Indicator				
First-Gen	299	11.9%	51	15.9%
Not First-Gen	2,212	88.1%	269	84.1%
Total Student Essays	2,511		320	

Note. SOC = Student of Color, W = White, EA = Early Action, RA = Regular Action, M = Male, F = Female

Essay Scoring & Methodology

Each selected essay was read and scored by one reader using the evaluation rubric described above (Figure 1). Initially, essays were assigned a score of 0-6 for each non-cognitive factor: 1) grammar and style; 2) grit; 3) intrinsic motivation; 4) creativity; 5) leadership; 6) community engagement; and, 7) cultural fluency. Scores were recorded in an Excel file that contained only the randomly generated ID number and a column for each non-cognitive factor score.

After coding the first few essays, challenges were encountered in applying the "no evidence" category on the rubric. For example, there is a difference between a student who writes an essay that does not touch upon any setbacks or hardship (grit), and a student who discusses setbacks or hardship, but fails to display grit in their response to those experiences. In response, the rubric was amended to include a "not applicable" (NA) category. A designation of NA indicated that the coder could not draw any conclusions regarding the prevalence of the non-cognitive factor based on the content presented in the essay. Any time the NA score applied, the letters "NA" were recorded in the database. Later, for data analysis purposes, the NA letters were deleted and the cells were left blank.

After essay coding was complete, the database was merged with the file from OIR containing relevant student data. Essay scores were calculated using four separate scoring methodologies. A detailed description of the different scoring methodologies is presented in Table 6. Because grammar and style are not non-cognitive factors, it was important to present an overall score that both included and excluded scores in this category. Overall essay scores 1a and 1b include a scoring designation of "not applicable" (NA) for each

46

rubric category and do not penalize a student for failing to address one of the six non-cognitive factors. 1a includes the grammar/style category and 1b omits grammar/style. Grammar/style could be evaluated for every essay in the sample, but not every essay displayed evidence of the six non-cognitive factors. Grammar/style scores had the highest prevalence (100%) over any other factor, and therefore had an inflating effect on the overall score. For that reason, it was especially important to retain separate scoring methodologies which include and exclude the grammar/style score.

Table 6

Summary of Essay Scoring Techniques and Methodology

Scoring System	Scoring Technique	Description of Scoring Methodology
1a	Score of 0: Based on the content presented in the essay, a conclusion can be drawn that the student displays no evidence of the non-cognitive factor. Score of "NA": Based on the content presented in the essay, no conclusions can be drawn regarding the prevalence of the non-cognitive factor.	Essay receives an overall essay score that presents an average of the grammar/style and non-cognitive scores with the denominator adjusted to reflect the number of categories (1-7) that received scores.
1b	Score of 0: Based on the content presented in the essay, a conclusion can be drawn that the student displays no evidence of the non-cognitive factor. Score of "NA": Based on the content presented in the essay, no conclusions can be drawn regarding the prevalence of the non-cognitive factor.	Essay receives an overall essay score that presents an average of the non-cognitive scores with the denominator adjusted to reflect the number of categories (1-6) that received scores. The grammar/style score is excluded from this calculation.
2a	Score of 0: Based on the content presented in the essay, a conclusion can be drawn that the student displays no evidence of the non-cognitive factor, OR no conclusions can be drawn regarding the prevalence of the non-cognitive factor. The "NA" designation does not exist in this scoring technique.	Essay receives an overall essay score that presents an average of the grammar/style and non-cognitive scores. The denominator is held constant and reflects the number of total scorable categories (7).
2b	Score of 0: Based on the content presented in the essay, a conclusion can be drawn that the student displays no evidence of the non-cognitive factor, OR no conclusions can be drawn regarding the prevalence of the non-cognitive factor. The "NA" designation does not exist in this scoring technique.	Essay receives an overall essay score that presents an average of the non-cognitive scores. The denominator is held constant and reflects the number of total scorable categories (6). Grammar/style score is excluded from this calculation.

Essay scores 2a and 2b do not include a scoring designation of NA. Instead, students received a score of zero if they failed to address one of the non-cognitive factors in their writing. 2a includes the grammar/style category and 2b omits grammar/style. The result was an artificially depressed overall score that penalizes students for not addressing all of the non-cognitive factors. The presence of non-cognitive factors in the student essay was highly dependent on the topic the student chose to write about, making it unrealistic that an essay would address all factors. Descriptive statistics for the four overall scores are presented in Table 7. As expected, based on the methodology, essay score 1a displays the highest mean ($M = 3.21$) followed by essay score 1b ($M = 1.99$), 2a ($M = 1.12$) and 2b ($M = 0.50$). With no evidence to suggest that any non-cognitive factor should carry more weight than another, and recognizing the constraints essay content places on the expression of non-cognitive traits, scoring methodology 1a and 1b present as the most reasonable scoring techniques. Overall essay scores using 2a and 2b methodology will not be presented in future results.

Table 7
Summary of Overall Essay Scores by Scoring Methodology

	Mean	Median	*SD*	Min	Max	*n*
Essay Score 1a	3.21	3.00	1.16	0.50	6.00	319
Essay Score 1b	1.99	2.00	1.31	0.00	6.00	319
Essay Score 2a	1.12	1.00	0.49	0.14	3.14	319
Essay Score 2b	0.50	0.43	0.41	0.00	2.43	319

Note. A detailed description of scoring techniques and methodologies are presented in Table 6.
1a: Includes the scoring designation of "NA," retains grammar/style score
1b: Includes the scoring designation of "NA," excludes grammar/style score
2a: Does not include a scoring designation of "NA," retains grammar/style score
2b: Does not include a scoring designation of "NA," excludes grammar/style score

Data Analysis

Descriptive statistics were calculated for pre-admission factors (high school GPA, ACT) and college outcomes (first-fall GPA) as well as for overall essay scores. Descriptive statistics were also calculated for the same variables disaggregated by applicant characteristics (e.g., early action vs. regular action). Not all essays addressed each non-cognitive factor, therefore the percentage of essays addressing each non-cognitive factor were calculated. Percentages were disaggregated by applicant characteristics and two sample t-tests were run between groups. Mean differences for overall and non-cognitive factor scores were calculated for each student characteristic and two sample t-tests were run between groups.

Correlation coefficients were calculated between overall scores, non-cognitive factor scores, pre-admission factors, and college outcomes. Correlation coefficients were calculated for the same variables disaggregated by applicant characteristic.

Findings

The study's findings are organized into three sections: 1) pre-admission factors and college outcomes; 2) operationalizing non-cognitive factors in essay evaluation; and 3) predicting pre-admission factors and college outcomes.

The first section provides descriptive information on academic achievement for the sample. Descriptive statistics disaggregated by applicant characteristic are provided for high school GPA, ACT, and first-fall GPA. The second section presents findings related to the operationalization of the non-cognitive factors in essay review. Descriptive statistics for overall essay scores and non-cognitive factor scores are presented as well as

mean differences between applicant characteristics. The third section correlates overall essay scores and non-cognitive factor scores with pre-admission factors and college outcomes.

Pre-Admission Factors and College Outcomes

Descriptive statistics for GPA, ACT, and first-fall college GPA for the sample are presented in Table 8. The overall mean high school GPA for the sample was 3.61 on a 4.0 scale. Descriptively, first-generation students displayed the highest mean high school GPA ($M = 3.69$) followed by females ($M = 3.67$) and early action applicants ($M = 3.63$). Males displayed the lowest mean high school GPA ($M = 3.55$). The mean overall ACT score was 26.78 (the ACT is scored out of 36). Descriptively, students who do not qualify as first-generation displayed the highest mean ACT ($M = 27.24$) followed by White students ($M = 27.06$) and male students ($M = 27.05$). First-generation students displayed the lowest mean ACT ($M = 24.33$). The overall mean first-fall college GPA was 3.07 on a 4.0 scale. Descriptively, White students displayed the highest mean first-fall GPA ($M = 3.14$) followed by females ($M = 3.13$) and early action applicants ($M = 3.12$). First-generation students displayed the lowest mean first-fall GPA ($M = 2.89$).

Table 8

Summary of Pre-Admission Factors and College Outcomes Disaggregated by Applicant Characteristic

	Mean	Median	*SD*	Min	Max	*n*
HS GPA	3.61	3.67	0.35	2.27	4	320
Early Action	3.63	3.67	0.34	2.27	4	167
Regular Action	3.60	3.67	0.36	2.47	4	153
Female	3.67	3.75	0.31	2.56	4	170
Male	3.55	3.62	0.38	2.27	4	150
White	3.63	3.69	0.34	2.27	4	224
Student of Color	3.58	3.64	0.36	2.47	4	96
First-Gen	3.69	3.79	0.29	3.04	4	51
Not First-Gen	3.60	3.65	0.35	2.27	4	269
ACT	26.78	26.50	3.52	16.00	35	320
Early Action	26.75	26.00	3.56	16.00	35	167
Regular Action	26.80	27.00	3.49	19.00	34	153
Female	26.53	26.00	3.42	16.00	34	170
Male	27.05	27.00	3.62	19.00	35	150
White	27.06	27.00	3.39	16.00	35	224
Student of Color	26.10	26.00	3.76	19.00	34	96
First-Gen	24.33	24.00	2.99	19.00	30	51
Not First-Gen	27.24	27.00	3.43	16.00	35	269
First-Fall GPA	3.07	3.23	0.80	0.00	4	318
Early Action	3.12	3.21	0.71	0.00	4	166
Regular Action	3.02	3.27	0.88	0.00	4	152
Female	3.13	3.38	0.85	0.00	4	170
Male	3.00	3.09	0.73	0.00	4	148
White	3.14	3.32	0.74	0.00	4	222
Student of Color	2.92	3.04	0.91	0.00	4	96
First-Gen	2.89	3.04	0.89	0.00	4	50
Not First-Gen	3.11	3.31	0.77	0.00	4	268

Note. Reported SAT scores were converted to ACT scores to create a universal standardized test score. If a student reported both SAT and ACT scores, the higher of the two scores was used.
HS GPA = High School GPA; First-Fall GPA = First semester grade point average at UVM

Operationalizing Non-Cognitive Factors in Essay Evaluation

Frequency of non-cognitive factors. The prevalence of non-cognitive factors varied greatly by essay topic and content. Therefore, not every essay addressed each of the non-cognitive factors and some non-cognitive factors were addressed more frequently than others. For example, every essay could be evaluated for grammar/style, but not every essay addressed the concept of grit. Table 9 presents the frequencies with which non-cognitive factors were addressed within the overall sample and by applicant characteristic. Column 1 displays the total number of essays (*n)* addressing each of the non-cognitive factors in the overall sample. In determining the count for each category, scores of zero were included. A score of zero indicated that the essay addressed the non-cognitive factor, but failed to meet the criteria for credit. Scores of zero were uncommon, appearing in fewer than 25 of the essays in the sample. Any score above zero indicated a positive expression of the non-cognitive factor and was also included in the count. Column 2 presents the percentage (%) of applications in the overall sample addressing each non-cognitive factor. Columns 3-14 display the prevalence of non-cognitive factors by applicant characteristic (e.g., early action vs. regular action) as well as the differences between groups. Two sample t-tests were run to determine whether the differences between the groups were significant.

Overall, grit was the most common non-cognitive factor addressed, appearing in 52.7% of essays in the sample. Intrinsic motivation was the next most common factor (31.0%) followed by creativity (26.6%), leadership (18.8%), cultural fluency (17.6%), and community engagement (14.1%).

Prevalence of non-cognitive factors differed across applicant characteristic groups, but two significant differences emerged. Essays written by White students were more likely to address community engagement than essays written by students of color (p <0.05). Additionally, essays written by students of color were more likely to address cultural fluency than essays written by White students (p <0.05).[4]

---

[4] It is important to remember that essay topic and content constrained the presentation of non-cognitive factors. Therefore, it would be inappropriate to conclude that White students are more engaged in their community than students of color. Instead, the findings display the percentage of essays addressing community engagement and how this differed by race/ethnicity. Similarly, it cannot be concluded that students of color are more culturally fluent than White students, but rather a higher percentage of essays written by students of color addressed the concept of cultural fluency.

Table 9

Percentage of Essays Mentioning Non-Cognitive Traits by Applicant Characteristic

	Overall		%			%			%			%		
	n	%	EA (*n* = 167)	RA (*n* = 152)	Δ	Female (*n* = 169)	Male (*n* = 150)	Δ	White (*n* = 223)	SOC (*n* = 96)	Δ	First-Gen (*n* = 51)	Not First-Gen (*n* = 268)	Δ
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Grammar/Style	319	100.0%	100.0%	100.0%	0.0%	100.0%	100.0%	0.0%	100.0%	100.0%	0.0%	100.0%	100.0%	0.0%
Grit	168	52.7%	53.9%	51.3%	2.6%	53.8%	51.3%	2.5%	54.3%	49.0%	5.3%	60.8%	51.1%	9.7%
Intrinsic Motivation	99	31.0%	31.1%	30.9%	0.2%	27.2%	35.3%	8.1%	34.1%	24.0%	10.1%	31.4%	31.0%	0.4%
Creativity	85	26.6%	31.1%	21.7%	9.4%	27.8%	25.3%	2.5%	29.6%	19.8%	9.8%	15.7%	28.7%	13.0%
Leadership	60	18.8%	22.2%	15.1%	7.0%	16.0%	22.0%	6.0%	20.6%	14.6%	6.0%	17.6%	19.0%	1.4%
Community Engagement	45	14.1%	17.4%	10.5%	6.8%	13.6%	14.7%	1.1%	17.0%	7.3%	9.7%*	9.8%	14.9%	5.1%
Cultural Fluency	56	17.6%	16.8%	18.4%	1.7%	20.7%	14.0%	6.7%	13.9%	26.0%	12.1%*	19.6%	17.2%	2.4%

Note. % = percentage of essays mentioning the non-cognitive factor within the respective category (e.g., percentage of early application essays mentioning grit). *n* = the total number of essays mentioning the non-cognitive factor.

SOC = student of color

p <0.05

Overall and non-cognitive factor scores. Differences in overall essay scores by applicant characteristic are presented in Tables 10 and 11. When grammar/style scores were included (Table 10), students of color had the highest mean overall score ($M = 3.53$) followed by females ($M = 3.33$) and first-generation students ($M = 3.27$). When grammar/style scores were omitted (Table 11), the highest mean overall scores were for first-generation ($M = 2.07$), early action applicants ($M = 2.05$), and students of color ($M = 2.03$). The only difference between scoring methodology 1a and 1b is the inclusion of the grammar/style score, therefore the differences in scores emphasize the variable effect grammar/style has within applicant groups.

Table 10
Overall Essay Scores by Applicant Characteristic (includes grammar/style score)

	Mean	Median	*SD*	Min	Max	*n*
Overall Essay Score 1a	3.21	3.00	1.16	0.50	6.00	319
Early Action	3.18	3.00	1.16	0.50	6.00	167
Regular Action	3.25	3.00	1.16	0.67	6.00	152
Female	3.33	3.00	1.25	0.57	6.00	169
Male	3.08	3.00	1.03	0.50	6.00	150
White	3.08	3.00	1.10	0.50	6.00	223
Student of Color	3.53	3.00	1.23	1.00	6.00	96
First-Gen	3.27	2.00	0.84	1.67	5.25	51
Not First-Gen	3.20	3.00	1.23	1.00	6.00	268

Table 11

Overall Essay Scores by Applicant Characteristic (excludes grammar/style score)

	Mean	Median	*SD*	Min	Max	*n*
Overall Essay Score 1b	1.99	2.00	1.31	0.00	6.00	319
Early Action	2.05	2.00	1.32	0.00	6.00	167
Regular Action	1.93	2.00	1.31	0.00	5.33	152
Female	1.97	2.00	1.42	0.00	6.00	169
Male	2.02	2.00	1.18	0.00	5.00	150
White	1.98	2.00	1.26	0.00	5.00	223
Student of Color	2.03	2.00	1.42	0.00	6.00	96
First-Gen	2.07	2.00	1.23	0.00	5.33	51
Not First-Gen	1.98	2.00	1.33	0.00	6.00	268

Table 12 presents the mean differences in overall essay scores and non-cognitive factor scores by applicant group. Overall essay scores excluding grammar/style (essay score 1b) did not differ significantly by application type, gender, race/ethnicity, or first-generation status. However, significant differences in overall essay scores including grammar/style (essay score 1a) were observed between White students ($M = 3.08$) and students of color ($M = 3.54$) with students of color displaying significantly higher mean overall essay scores than White students ($p <0.01$). This further suggests that grammar/style scores play an important role in overall essay scoring. When non-cognitive factor scores were disaggregated by applicant characteristic, significant differences in mean scores were observed in grammar/style, creativity, and cultural fluency. Students of color ($M = 4.73$) displayed significantly higher average grammar/style scores than White students ($M = 4.22$, $p < 0.001$). Higher average creativity scores were observed in the regular action applicant essays ($M = 2.06$) compared with early action applicant essays ($M = 1.44$, $p < 0.05$). Additionally, students of color displayed higher mean scores in cultural fluency ($M = 2.88$) than White students ($M = 2.03$, $p < 0.05$).

Table 12

Mean Differences in Overall Score and Non-Cognitive Factor Score by Applicant Characteristic

	Mean			Mean			Mean			Mean		
	EA	RA	Δ	Female	Male	Δ	White	SOC	Δ	First-Gen	Not First-Gen	Δ
Overall Essay Score 1a	3.18	3.25	0.07	3.33	3.08	0.25	3.08	3.54	0.46**	3.27	3.21	0.06
Overall Essay Score 1b	2.05	1.93	0.12	1.97	2.02	0.06	1.98	2.03	0.05	2.07	1.98	0.09
Grammar/Style	4.34	4.38	0.04	4.44	4.27	0.18	4.20	4.73	0.53***	4.59	4.32	0.27
Grit	2.42	2.17	0.26	2.31	2.30	0.01	2.20	2.57	0.38	2.52	2.26	0.26
Intrinsic Motivation	2.44	2.32	0.12	2.50	2.28	0.22	2.43	2.22	0.22	1.88	2.48	0.61
Creativity	1.44	2.06	0.62*	1.64	1.74	0.10	1.67	1.74	0.07	2.25	1.62	0.63
Leadership	1.92	2.13	0.21	2.00	2.00	0.00	1.96	2.14	0.19	2.44	1.92	0.52
Community Engagement	1.86	2.44	0.58	1.83	2.32	0.49	2.05	2.14	0.09	2.60	2.00	0.60
Cultural Fluency	2.43	2.39	0.04	2.34	2.52	0.18	2.03	2.88	0.85*	2.50	2.39	0.11

Note. Overall Essay Score 1a includes the grammar/style score. Overall Essay Score 1b excludes the grammar/style score. EA = Early Action; RA = Regular Action; SOC = Students of Color; First-Gen = First Generation College Student
*p<.05; **p<.01; ***p<.001

58

Predicting Pre-Admission Factors and College Outcomes

Correlation coefficients between non-cognitive factors, pre-application factors, and college outcomes are presented in Tables 13 and 14. Very weak correlations were found between overall essay scores and high school GPA, ACT, or first-fall GPA. Similarly, no weak correlations were observed between non-cognitive factor scores and high school GPA, ACT, or fist-fall GPA when the sample was analyzed as a whole (Table 13).

When each rubric factor is disaggregated by applicant characteristic, a few notable trends arose (Table 14). Intrinsic motivation is weakly correlated ($r = .25$) with high school GPA for first-generation students. Additionally, community engagement is moderately correlated ($r = .38$) with first-fall GPA for first-generation students. For students of color, community engagement appears negatively correlated with high school GPA ($r = -.68$) while cultural fluency is moderately correlated with first-fall GPA ($r = .32$).

Table 13
Correlations Between Overall Essay Scores, Non-Cognitive Factor Scores, and Pre-Admission Factors/College Outcomes

	HS GPA	ACT	First-Fall GPA
Overall Essay Score 1a	-0.05	-0.06	-0.05
Overall Essay Score 1b	0.01	0.00	-0.14
Grammar	-0.08	0.02	0.02
Grit	-0.07	-0.13	-0.19
Intrinsic Motivation	-0.11	-0.13	-0.21
Creativity	-0.05	-0.14	-0.20
Leadership	0.00	-0.04	-0.15
Community Engagement	-0.10	0.12	-0.13
Cultural Fluency	0.14	0.03	-0.01

Table 14

Correlations Between Non-Cognitive Factor Scores and Pre-Admission Factors/College Outcomes Disaggregated by Applicant Characteristic

	High School GPA	ACT Score	First-Fall GPA
Grammar/Style	-0.08	0.02	0.00
EA	-0.07	0.01	-0.08
RA	-0.08	0.02	0.09
Female	-0.04	0.04	0.13
Male	-0.14	0.00	-0.17
White	-0.09	0.02	-0.06
SOC	-0.02	0.08	0.21
First-Gen	0.12	0.10	0.32
Not First-Gen	-0.11	0.03	-0.03
Grit	-0.07	-0.13	-0.19
EA	-0.07	-0.17	-0.14
RA	-0.07	-0.07	-0.25
Female	-0.13	-0.03	-0.22
Male	-0.01	-0.26	-0.13
White	-0.08	-0.06	-0.18
SOC	0.00	-0.24	-0.17
First-Gen	0.00	0.02	-0.15
Not First-Gen	-0.09	-0.14	-0.18
Intrinsic Motivation	-0.11	-0.13	-0.21
EA	-0.04	-0.03	-0.24
RA	-0.22	-0.32	-0.19
Female	0.00	-0.13	-0.09
Male	-0.20	-0.11	-0.32
White	-0.11	-0.24	-0.24
SOC	-0.19	0.12	-0.19
First-Gen	0.25	0.02	0.17
Not First-Gen	-0.11	-0.23	-0.25
Creativity	-0.05	-0.14	-0.20
EA	0.04	-0.02	-0.19
RA	-0.07	-0.27	-0.07
Female	-0.01	-0.19	-0.17
Male	-0.06	-0.09	-0.25
SOC	0.01	-0.38	-0.12
White	-0.07	-0.04	-0.24
First-Gen	-0.05	-0.20	-0.03
Not First-Gen	-0.07	-0.09	-0.21

Correlations Between Non-Cognitive Factor Scores and Pre-Admission Factors/College Outcomes Disaggregated by Applicant Characteristic, Continued.

	High School GPA	ACT Score	First-Fall GPA
Leadership	0.00	-0.04	-0.15
EA	-0.03	0.10	-0.15
RA	0.03	-0.24	-0.13
Female	-0.13	-0.07	-0.20
Male	0.10	-0.01	-0.07
White	0.02	0.04	-0.08
SOC	-0.12	-0.23	-0.32
First-Gen	0.03	-0.04	-0.21
Not First-Gen	-0.04	0.01	-0.14
Community Engagement	-0.10	0.12	-0.13
EA	-0.20	0.06	-0.29
RA	-0.05	0.29	0.12
Female	-0.17	0.16	-0.04
White	-0.03	0.15	-0.13
Male	0.03	0.05	-0.19
SOC	-0.68	0.03	-0.24
First-Gen	0.07	0.00	0.38
Not First-Gen	-0.12	0.19	-0.16
Cultural Fluency	0.14	0.03	-0.01
EA	0.33	-0.01	-0.07
RA	-0.02	0.06	0.02
Female	0.22	0.08	-0.03
Male	0.08	-0.09	0.08
White	-0.01	-0.05	-0.20
SOC	0.26	0.19	0.32
First-Gen	-0.53	-0.21	-0.03
Not First-Gen	0.19	0.07	0.00

Summary

The rubric developed and tested in this study was able to detect the presence and strength of non-cognitive factors in student essays, but its ability to do so was highly dependent on essay topic and content. The tool was not correlated with pre-admission factors or first semester GPA. However, these results do not necessarily render the rubric void of value.

Discussion

The practice of holistic admissions involves the review of multiple elements of a student's academic history and personal context. Holistic admissions encourages individualized review and strives to deemphasize the importance of any single component of the application (Bastedo et al., 2016; Lucido, 2014). It is therefore noteworthy that this study focused solely on the essay as a vehicle through which to review applicant characteristics and predict college outcomes. This is a narrow window and provides only a 650-word snapshot into a student's life and personal experience. This limitation was evident in an overview of the percentage of essays addressing each non-cognitive factor (Table 9). The overall prevalence of each non-cognitive factor supported the idea that non-cognitive factors can be operationalized in essay review. However, some non-cognitive traits were more frequently addressed than others, highlighting the challenges of using the essay as a sole mode of evaluation. Grit appears to be the most common non-cognitive factor addressed, appearing in just over half of the essays, but community engagement was addressed in only 14.1% of the essays. The ability to score each factor was heavily dependent on the topic and content of the essay, suggesting that some non-cognitive factors may be better evaluated in other areas of the application. This

concern was voiced during the interviews with admissions experts. Interestingly, when asked which non-cognitive factors might be the least topic-dependent, four of the five interviewees concluded grit and intrinsic motivation. This prediction was substantiated by the fact that grit and intrinsic motivation had the highest representation in the sample, appearing in 52.7% and 31.0% of the essays, respectively.

Despite presentation of non-cognitive factors being topic-dependent, this study offers insights into the value of operationalizing essay review. The essays were written in response to a range of standardized prompts. Some prompts encouraged students to describe a background, identity, interest, or talent that is meaningful to their application. Others asked students to describe a formative event that prompted growth or a time they experienced challenges. A unifying theme in the essay prompts is that they promote reflection and disclosure of a meaningful experience. Theoretically, the content of a student's essay reflects an experience that was truly impactful. Under this assumption, understanding the prevalence of non-cognitive traits and how they vary by applicant group becomes important. In this study, essays written by first-generation students were more likely to display evidence of grit (60.8%) than essays written by students who were not first-generation (51.1%). The small sample size made it impossible to determine whether this difference was statistically significant, but it is noteworthy that the group of first-generation students in the sample also had the lowest mean ACT score. This finding provides evidence in support of holistic review and displays the ways in which institutional values may converge through operationalized review of non-cognitive factors in the student essay. Concerns associated with weaker academic credentials may be assuaged by the presentation of valued non-cognitive traits in the essay.

It is important not to overgeneralize the presentation of non-cognitive factors by applicant characteristic. For example, students of color were less likely than White students to provide examples of community engagement in their applicant essays. However, because essay content constrains the presentation of non-cognitive factors, it would be unreasonable to conclude that students of color do not engage in their communities. This highlights the limitations of using the same rubric to evaluate essays addressing different topics. A tailored supplemental essay prompt asking students to specifically discuss the ways they have engaged with the community may be more effective. Additionally, the essay simply may not be the best mode through which to evaluate community engagement. Instead, more accurate evaluations may occur through review of the extracurricular grid or letters of recommendation. The same could be said for all of the non-cognitive factors. A valuable extension of this study would be to apply the rubric to the entire college application, not just the essay.

The most significant findings regarding the ability to operationalize non-cognitive factors are evident in Table 12. A mean difference comparison between applicant characteristics for overall scores and non-cognitive factor scores revealed that students of color scored significantly higher than White students cultural fluency, grammar/style, and overall essay scores (when grammar/style was included). These findings not only assist in validating the rubric as an evaluation tool, but also have important implications for equity. In this sample, students of color displayed slightly lower average high school GPA's and ACT scores than White students. Given that standardized test scores have been criticized for potentially displaying ethnic and cultural bias (Freedle, 2003;

Santelices & Wilson, 2010), it's exciting to see that a rubric prioritizing non-cognitive factors does not perpetuate the same inequalities.

A limitation of the sample is that every essay represents a student who was first admitted and then matriculated to UVM. The students are only in their second semester at UVM, therefore there are few academic outcomes with which essay scores could be correlated. First-fall GPA only reveals a small component of the overall student experience and isn't necessarily an indicator of whether a student is thriving on UVM's campus. An extension of this study would be to track the same population over the next three years and reevaluate the predictive nature of the rubric when other outcome measures (retention, 4-year cumulative GPA) become available. It is also important to evaluate the definition of success.

While this particular essay evaluation rubric does not correlate with first-fall GPA, that does not nullify the value of expressing the non-cognitive traits. Every admission expert interviewed during the rubric development process agreed that the non-cognitive factors UVM was focused on reflected characteristics they would value in an applicant or alumnus of their own institution. An anecdotal analysis of one essay offers insight into why essay scores may not directly correlate with college grades. One student wrote an essay highlighting their passion for entrepreneurship. The student described a business they started and their goals for future endeavors. The student displayed a high level of intrinsic motivation in their pursuits, but received scores of "NA" in every other non-cognitive category. It is likely the student experienced setbacks and needed to engage in the community to market and promote their business, but if these experiences occurred, they were not discussed. Additionally, the essay contained a number of

grammatical errors and wasn't organized effectively, producing a lower grammar/style score. The result was a low overall essay score for a student displaying strong intrinsic motivation and a clear goal of enrolling and graduating from UVM's College of Business. If this same student were to graduate from UVM with a low GPA, but go on to start a successful business, would that be a successful outcome? A non-cognitive trait can be positive, desirable, and still not correlate with college performance. Akos & Kretchmar (2016) draw a similar conclusion in their study of grit as a non-cognitive predictor of college outcomes.

Implications

What implications do these findings have for how student essays are used in college admissions decisions? One clear question is left unanswered: What is the purpose of the college admissions essay? Essay topics range significantly making it difficult to standardize review. Admissions offices need to decide how they choose to use the essay in their review process. If it is to be used as a tool for measuring specific outcomes, the outcomes should be pre-defined and a supplemental essay topic could be designed to address the specific question.

This study provides evidence that non-cognitive traits can be operationalized for review in the college admissions essay. The study presents a tool that assists in quantifying evidence of 1) grit; 2) intrinsic motivation; 3) creativity; 4) leadership; 5) community engagement; and 6) cultural fluency in the application essay. The tool requires further testing on a larger sample with increased college outcome measures (e.g., 4-year cumulative GPA, retention data). It also requires testing for inter-rater reliability

and should be used at multiple institutions before any conclusions regarding its predictive validity can be drawn. However, there is a critical need for standardized assessment tools within the field of college admissions and this instrument takes a first, promising step towards the direction of operationalizing and prioritizing non-cognitive traits in the admissions process.

Bibliography

Akos, P., & Kretchmar, J. (2016). Investigating Grit at a Non-Cognitive Predictor of College Success. *The Review of Higher Education*, *40*(2), 163–186. https://doi.org/10.1353/rhe.2017.0000

Albanese, M. A., Snow, M. H., Skochelak, S. E., Huggett, K. N., & Farrell, P. M. (2003). Assessing Personal Qualities in Medical School Admissions. *Academic Medicine*, *78*(3), 313.

Astin, A. W., & Sax, L. J. (1998). How Undergraduates Are Affected by Service Participation. *Journal of College Student Development*, *39*(3), 251–263.

Atkinson, R. (2001). Standardized Tests and Access to American Universities. Retrieved from https://escholarship.org/uc/item/6182126z

Baird, L. L. (1977). What Graduate and Professional School Students Think About Admissions Tests. NCME Measurement in Education. A Series of Special Reports of the National Council on Measurement in Education. *NCME*, *7*(3). Retrieved from http://eric.ed.gov/?id=ED157903

Balogun, J. A., Karacoloff, L. A., & Farina, N. T. (1986). Predictors of academic achievement in physical therapy. *Physical Therapy*, *66*(6), 976–980.

Barron, J. M., Ewing, B. T., & Waddell, G. R. (2000). The Effects of High School Athletic Participation on Education and Labor Market Outcomes. *The Review of Economics and Statistics*, *82*(3), 409–421. https://doi.org/10.1162/003465300558902

Bastedo, M. N., Howard, J. E., & Flaster, A. (2016). Holistic Admissions After Affirmative Action: Does "Maximizing" the High School Curriculum Matter? *Educational Evaluation and Policy Analysis*, *38*(2), 389–409. https://doi.org/10.3102/0162373716642407

Beck, S. W., & Jeffery, J. V. (2007). Genres of high-stakes writing assessments and the construct of writing competence. *Assessing Writing*, *12*(1), 60–79. https://doi.org/10.1016/j.asw.2007.05.001

Berchulc, C. M., Wade, G. A., & Seidner, K. K. (1987). Predicting Academic Achievement of Undergraduate Occupational Therapy Students: Preliminary Results. *The Occupational Therapy Journal of Research*, *7*(4), 245–248. https://doi.org/10.1177/153944928700700406

Bowman, N., Brandenberger, J., Lapsley, D., Hill, P., & Quaranto, J. (2010). Serving in College, Flourishing in Adulthood: Does Community Engagement During the College Years Predict Adult Well-Being? *Applied Psychology: Health and Well-Being*, *2*(1), 14–34. https://doi.org/10.1111/j.1758-0854.2009.01020.x

Breland, H. M., Bridgeman, B., & Fowles, M. E. (1999). Writing Assessment in Admission to Higher Education: Review and Framework. *ETS Research Report Series*, *1999*(1), i–37. https://doi.org/10.1002/j.2333-8504.1999.tb01801.x

Breland, H. M., & Jones, R. J. (1988). Remote Scoring of Essays. *ETS Research Report Series*, *1988*(1), i–39. https://doi.org/10.1002/j.2330-8516.1988.tb00260.x

Brereton, J. (1996). *The Origins of Composition Studies in the American College, 1875–1925*. University of Pittsburgh Press.

Bridgeman, B., & Lewis, C. (1994). The Relationship of Essay and Multiple-Choice Scores With Grades in College Courses. *Journal of Educational Measurement*, *31*(1), 37–50. https://doi.org/10.1111/j.1745-3984.1994.tb00433.x

Carlone, H. B. (2004). The cultural production of science in reform-based physics: Girls' access, participation, and resistance. *Journal of Research in Science Teaching*, *41*(4), 392–414. https://doi.org/10.1002/tea.20006

Chang, M. J. (2001). The Positive Educational Effects of Racial Diversity on Campus. In G. Orfield, *Diversity Challenged: Evidence on the Impact of Affirmative Action*. Retrieved from http://eric.ed.gov/?id=ED456198

Clinedinst, M., & Koranteng, A. M. (2017). *State of College Admission* (Report). National Association for College Admission Counseling (NACAC). Retrieved from https://vtechworks.lib.vt.edu/handle/10919/83148

Clinedinst, M., & Patel, P. (2018). *State of College Admission*. National Association for College Admission Counseling (NACAC).

Collins, A., & Gentner, D. (1980). A framework for a cognitive theory of writing. *Cognitive Processes in Writing*, 51–72.

Cotter, K. N., Pretz, J. E., & Kaufman, J. C. (2016). Applicant Extracurricular Involvement Predicts Creativity Better Than Traditional Admissions Factors. *Psychology of Aesthetics, Creativity, and the Arts*, *10*(1), 2–13. https://doi.org/10.1037/a0039831

Cress, C. M., Astin, H. S., Zimmerman-Oster, K., & Burkhardt, J. C. (2001). Developmental outcomes of college students' involvement in leadership activities. *Journal of College Student Development*, *42*(1), 15–27.

Cropley, D. H. (2015). Promoting creativity and innovation in engineering education. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(2), 161–171. https://doi.org/10.1037/aca0000008

Crosby, F. J., Iyer, A., Clayton, S., & Downing, R. A. (2003). Affirmative action: Psychological data and the policy debates. *American Psychologist*, *58*(2), 93–115. https://doi.org/10.1037/0003-066X.58.2.93

Cumming, A., Kantor, R., Baba, K., Erdosy, U., Eouanzoui, K., & James, M. (2005). Differences in written discourse in independent and integrated prototype tasks for next generation TOEFL. *Assessing Writing*, *10*(1), 5–43. https://doi.org/10.1016/j.asw.2005.02.001

Diederich, P. B. (1974). Measuring Growth in English. Retrieved from http://eric.ed.gov/?id=ED097702

Dollinger, S. J. (2011). "Standardized minds" or individuality? Admissions tests and creativity revisited. *Psychology of Aesthetics, Creativity, and the Arts*, *5*(4), 329–341. https://doi.org/10.1037/a0023659

Duckworth, A. (2016). *Grit: The Power of Passion and Perseverance* (1 edition). New York: Scribner.

Duckworth, A., & Gross, J. J. (2014). Self-Control and Grit: Related but Separable Determinants of Success. *Current Directions in Psychological Science*, *23*(5), 319–325. https://doi.org/10.1177/0963721414541462

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92*(6), 1087–1101. https://doi.org/10.1037/0022-3514.92.6.1087

Duckworth, A., & Quinn, P. D. (2009). Development and Validation of the Short Grit Scale (Grit–S). *Journal of Personality Assessment*, *91*(2), 166–174. https://doi.org/10.1080/00223890802634290

Dumais, S. A. (2008). Cohort and Gender Differences in Extracurricular Participation: The Relationship Between Activities, Math Achievement, and College Expectations. *Sociological Spectrum*, *29*(1), 72–100. https://doi.org/10.1080/02732170802480543

Eccles, J. S., & Barber, B. L. (1999). Student Council, Volunteering, Basketball, or Marching Band: What Kind of Extracurricular Involvement Matters? *Journal of Adolescent Research*, *14*(1), 10–43. https://doi.org/10.1177/0743558499141003

Eitle, T. M. (2005). Do Gender and Race Matter? Explaining the Relationship Between Sports Participation and Achievement. *Sociological Spectrum*, *25*(2), 177–195. https://doi.org/10.1080/02732170590883997

Enright, M. K., & Gitomer, D. (1989). Toward a Description of Successful Graduate Students. *ETS Research Report Series*, *1989*(1), i–37. https://doi.org/10.1002/j.2330-8516.1989.tb00335.x

Feist, G. J. (1998). A Meta-Analysis of Personality in Scientific and Artistic Creativity. *Personality and Social Psychology Review*, *2*(4), 290–309. https://doi.org/10.1207/s15327957pspr0204_5

Feist, G. J., & Barron, F. X. (2003). Predicting creativity from early to late adulthood: Intellect, potential, and personality. *Journal of Research in Personality*, *37*(2), 62–88. https://doi.org/10.1016/S0092-6566(02)00536-6

First-Year Applicants. (n.d.). Retrieved February 15, 2019, from https://www.uvm.edu/admissions/undergraduate/first_year_applicants

Florida, R. (2002). *The rise of the creative class: and how it's transforming work, leisure, community and everyday life / Richard Florida.* New York, NY: Basic Books.

Flower, L., & Hayes, J. R. (1981). A Cognitive Process Theory of Writing. *College Composition and Communication*, *32*(4), 365–387. https://doi.org/10.2307/356600

Fouad, N. A., & Smith, P. L. (1996). A Test of a Social Cognitive Model for Middle School Students: Math and Science. *Journal of Counseling Psychology*, *43*(3), 338–346. https://doi.org/10.1037/0022-0167.43.3.338

Fouad, N. A., Smith, P. L., & Zao, K. E. (2002). Across Academic Domains: Extensions of the Social-Cognitive Career Model. *Journal of Counseling Psychology*, *49*(2), 164–171. https://doi.org/10.1037/0022-0167.49.2.164

Fredricks, J. A. (2012). Extracurricular Participation and Academic Outcomes: Testing the Over-Scheduling Hypothesis. *Journal of Youth and Adolescence*, *41*(3), 295–306. https://doi.org/10.1007/s10964-011-9704-0

Freedle, R. (2003). Correcting the SAT's Ethnic and Social-Class Bias: A Method for Reestimating SAT Scores. *Harvard Educational Review*, *73*(1), 1–43. https://doi.org/10.17763/haer.73.1.8465k88616hn4757

Galton, F. (1869). *Hereditary genius: An inquiry into its laws and consequences* (Vol. 27). Macmillan.

Gardner, M., Roth, J., & Brooks-Gunn, J. (2008). Adolescents' participation in organized activities and developmental success 2 and 8 years after high school: Do sponsorship, duration, and intensity matter? *Developmental Psychology*, *44*(3), 814–830. http://dx.doi.org.ezproxy.uvm.edu/10.1037/0012-1649.44.3.814

Gerber, S. B. (1996). Extracurricular activities and academic achievement. *Journal of Research & Development in Education*, *30*(1), 42–50.

Getzels, J. W., & Jackson, P. W. (1962). *Creativity and intelligence: Explorations with gifted students*. Oxford, England: Wiley.

GlenMaye, L., & Oakes, M. (2002). Assessing Suitability of MSW Applicants through Objective Scoring of Personal Statements. *Journal of Social Work Education*, *38*(1), 67–82. https://doi.org/10.1080/10437797.2002.10779083

Gose, B. (2005, February 25). Looking Beyond "U.S. News," Rethinking the SAT. *The Chronicle of Higher Education*. Retrieved from https://www.chronicle.com/article/Looking-Beyond-US-News/34869

Gough, H. G. (1992). Assessment of Creative Potential in Psychology and the Development of a Creative Temperament Scale for the CPI. *Advances in Psychological Assessment*, 225–257. https://doi.org/10.1007/978-1-4757-9101-3_8

Hart, D, Matsuba, M. K., & Atkins, R. (2008). The moral and civic effects of learning to serve. In *Handbook of Moral Character Education* (pp. 484–499).

Hart, Daniel, Donnelly, T. M., Youniss, J., & Atkins, R. (2007). High School Community Service as a Predictor of Adult Voting and Volunteering. *American Educational Research Journal*, *44*(1), 197–219. https://doi.org/10.3102/0002831206298173

Harvard Admissions Lawsuit. (n.d.). Retrieved February 15, 2019, from https://admissionscase.harvard.edu/home

Hazari, Z., Potvin, G., Tai, R. H., & Almarode, J. (2010). For the love of learning science: Connecting learning orientation and career productivity in physics and chemistry. *Physical Review Special Topics - Physics Education Research*, *6*(1), 010107. https://doi.org/10.1103/PhysRevSTPER.6.010107

Henderson-King, D., & Kaleta, A. (2000). Learning about Social Diversity: The Undergraduate Experience and Intergroup Tolerance. *Journal of Higher Education*, *71*(2), 142–64. https://doi.org/10.2307/2649246

Hoffman, J. L. (2002). The impact of student cocurricular involvement on student success: Racial and religious differences. *Journal of College Student Development; Baltimore*, *43*(5), 712.

Hunt, H. D. (2005). The Effect of Extracurricular Activities in the Educational Process: Influence on Academic Outcomes? *Sociological Spectrum*, *25*(4), 417–445. https://doi.org/10.1080/027321790947171

Huot, B. (1990). The Literature of Direct Writing Assessment: Major Concerns and Prevailing Trends. *Review of Educational Research*, *60*(2), 237–263. https://doi.org/10.3102/00346543060002237

Kaufman, J. C. (2006). Self-reported differences in creativity by ethnicity and gender. *Applied Cognitive Psychology*, *20*(8), 1065–1082. https://doi.org/10.1002/acp.1255

Kaufman, J. C., & Agars, M. D. (2009). Being creative with the predictors and criteria for success. *American Psychologist*, *64*(4), 280–281. https://doi.org/10.1037/a0014926

Kaufman, J. C., Agars, M. D., & Lopez-Wagner, M. C. (2008). The role of personality and motivation in predicting early college academic success in non-traditional students at a Hispanic-serving institution. *Learning and Individual Differences*, *18*(4), 492–496. https://doi.org/10.1016/j.lindif.2007.11.004

Kellogg, R. T., & Raulerson, B. A. (2007). Improving the writing skills of college students. *Psychonomic Bulletin & Review*, *14*(2), 237–242. https://doi.org/10.3758/BF03194058

Kirchner, G. L., & Holm, M. B. (1997). Prediction of Academic and Clinical Performance of Occupational Therapy Students in an Entry-Level Master's Program. *American Journal of Occupational Therapy*, *51*(9), 775–779. https://doi.org/10.5014/ajot.51.9.775

Koenig, T. W., Parrish, S. K., Terregino, C. A., Williams, J. P., Dunleavy, D. M., & Volsch, J. M. (2013). Core Personal Competencies Important to Entering Students' Success in Medical School: What Are They and How Could They Be Assessed Early in the Admission Process? *Academic Medicine*, *88*(5), 603–613. https://doi.org/10.1097/ACM.0b013e31828b3389

Kretchmar, J. (2006). Assessing the Reliability of Ratings Used in Undergraduate Admission Decisions. *Journal of College Admission*. Retrieved from https://eric.ed.gov/?id=EJ783935

Kuhn, P., & Weinberger, C. (2005). Leadership Skills and Wages. *Journal of Labor Economics*, *23*(3), 395–436. https://doi.org/10.1086/430282

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic Performance, Career Potential, Creativity, and Job Performance: Can One Construct Predict Them All? *Journal of Personality and Social Psychology*, *86*(1), 148–161. https://doi.org/10.1037/0022-3514.86.1.148

Lareau, A. (2011). *Unequal childhoods: Class, race, and family life.* University of California Press.

Lent, R. W., Brown, S. D., Schmidt, J., Brenner, B., Lyons, H., & Treistman, D. (2003). Relation of Contextual Supports and Barriers to Choice Behavior in Engineering Majors: Test of Alternative Social Cognitive Models. *Journal of Counseling Psychology*, *50*(4), 458–465. https://doi.org/10.1037/0022-0167.50.4.458

Lepper, M. R., Corpus, J. H., & Iyengar, S. S. (2005). Intrinsic and Extrinsic Motivational Orientations in the Classroom: Age Differences and Academic Correlates. *Journal of Educational Psychology*, *97*(2), 184–196. https://doi.org/10.1037/0022-0663.97.2.184

Lleras, C. (2008). Do skills and behaviors in high school matter? The contribution of noncognitive factors in explaining differences in educational attainment and earnings. *Social Science Research*, *37*(3), 888–902. https://doi.org/10.1016/j.ssresearch.2008.03.004

Lozano, F. A. (2008). Language, high school leadership and the postsecondary outcomes of Hispanic students. *Economics of Education Review*, *27*(3), 342–353. https://doi.org/10.1016/j.econedurev.2007.01.003

Lubinski, D., Benbow, C. P., Webb, R. M., & Bleske-Rechek, A. (2006). Tracking Exceptional Human Capital Over Two Decades. *Psychological Science*, *17*(3), 194–199. https://doi.org/10.1111/j.1467-9280.2006.01685.x

Lucido, J. A. (2014). How admission decisions get made. *Handbook of Strategic Enrollment Management*, 147–173.

Maddi, S. R., Matthews, M. D., Kelly, D. R., Villarreal, B., & White, M. (2012). The Role of Hardiness and Grit in Predicting Performance and Retention of USMA Cadets. *Military Psychology*, *24*(1), 19–28. https://doi.org/10.1080/08995605.2012.639672

Mamlet, R., & VanDeVelde, C. (2011). *College admission: From application to acceptance, step by step*. Three Rivers Press (CA).

Marsh, H., & Kleitman, S. (2002). Extracurricular School Activities: The Good, the Bad, and the Nonlinear. *Harvard Educational Review*, *72*(4), 464–515. https://doi.org/10.17763/haer.72.4.051388703v7v7736

Marsh, H. W. (1991). Employment During High School: Character Building or a Subversion of Academic Goals? *Sociology of Education*, *64*(3), 172–189. https://doi.org/10.2307/2112850

Matteucci, A. J., Park, A. D., Patrick, S. D., Galla, B. M., & Duckworth, A. L. (2016). *Proof of perseverance and passion for long-term goals: The development and validation of a biodata measure of grit.*

Mattson, C. E. (2007). Beyond Admission: Understanding Pre-College Variables and the Success of At-Risk Students. *Journal of College Admission*. Retrieved from https://eric-ed-gov.ezproxy.uvm.edu/?id=EJ783950

Mayer, R. E. (1999). *Fifty Years of Creativity Research*. Cambridge University Press.
McDonald, J. D. (2008). Measuring personality constructs: The advantages and disadvantages of self-reports, informant reports and behavioural assessments. *Enquire*, *1*(1), 1–19.

Metz, E., & Youniss, J. (2003). A Demonstration That School-Based Required Service Does Not Deter—But Heightens—Volunteerism. *PS: Political Science & Politics*, *36*(2), 281–286. https://doi.org/10.1017/S1049096503002221

Morrow-Howell, N., Hinterlong, J., Rozario, P. A., & Tang, F. (2003). Effects of Volunteering on the Well-Being of Older Adults. *The Journals of Gerontology: Series B*, *58*(3), S137–S145. https://doi.org/10.1093/geronb/58.3.S137

Niehaus, K., Rudasill, K. M., & Adelson, J. L. (2012). Self-Efficacy, Intrinsic Motivation, and Academic Outcomes Among Latino Middle School Students Participating in an After-School Program. *Hispanic Journal of Behavioral Sciences*, *34*(1), 118–136. https://doi.org/10.1177/0739986311424275

Niessen, A. S. M., Meijer, R. R., & Tendeiro, J. N. (2017). Measuring non-cognitive 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

Office of Institutional Research. (n.d.). Admissions Information. Retrieved February 15, 2019, from https://www.uvm.edu/~oir/?Page=adinfo.html&SM=submenu_adm_enroll.html

Palmer, S. R. (2001). A Policy Framework for Reconceptualizing the Legal Debate Concerning Affirmative Action in Higher Education. In G. Orfield, *Diversity*

challenged: Evidence on the impact of affirmative action (pp. 49–80). Retrieved from http://eric.ed.gov/?id=ED456193

Pascarella, E. T., Palmer, B., Moye, M., & Pierson, C. T. (2001). Do diversity experiences influence the development of critical thinking? *Journal of College Student Development*, *42*(3), 257–271.

Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When Small Words Foretell Academic Success: The Case of College Admissions Essays. *PLOS ONE*, *9*(12), e115844. https://doi.org/10.1371/journal.pone.0115844

Pike, G. R., Kuh, G. D., & Gonyea, R. M. (2007). Evaluating the Rationale for Affirmative Action in College Admissions: Direct and Indirect Relationships between Campus Diversity and Gains in Understanding Diverse Groups. *Journal of College Student Development*, *48*(2), 166–182. https://doi.org/10.1353/csd.2007.0018

Piliavin, J. A., & Siegl, E. (2007). Health Benefits of Volunteering in the Wisconsin Longitudinal Study. *Journal of Health and Social Behavior*, *48*(4), 450–464. https://doi.org/10.1177/002214650704800408

Powers, D. E., & Fowles, M. E. (1998). Effects of Preexamination Disclosure of Essay Topics. *Applied Measurement in Education*, *11*(2), 139–157. https://doi.org/10.1207/s15324818ame1102_2

Preiss, D. D., Castillo, J. C., Grigorenko, E. L., & Manzi, J. (2013). Argumentative writing and academic achievement: A longitudinal study. *Learning and Individual Differences*, *28*, 204–211. https://doi.org/10.1016/j.lindif.2012.12.013

Pretz, J. E., & Kaufman, J. C. (2017). Do Traditional Admissions Criteria Reflect Applicant Creativity? *The Journal of Creative Behavior*, *51*(3), 240–251. https://doi.org/10.1002/jocb.120

Puccio, G. J., & Cabra, J. F. (2010, August). Organizational Creativity. https://doi.org/10.1017/CBO9780511763205.011

Reeb, R. N., Katsuyama, R. M., Sammon, J. A., & Yoder, D. S. (1998). The Community Service Self-Efficacy Scale: Evidence of Reliability, Construct Validity, and Pragmatic Utility. *Michigan Journal of Community Service Learning*, *5*, 48–57.

Robertson-Kraft, C., & Duckworth, A. L. (2014). True Grit: Trait-level Perseverance and Passion for Long-term Goals Predicts Effectiveness and Retention among Novice Teachers. *Teachers College Record (1970)*, *116*(3). Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4211426/

Robinson, R. L., Navea, R., & Ickes, W. (2013). Predicting Final Course Performance From Students' Written Self-Introductions: A LIWC Analysis. *Journal of Language and Social Psychology*, *32*(4), 469–479. https://doi.org/10.1177/0261927X13476869

Ryan, R., & Deci, E. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. http://dx.doi.org.ezproxy.uvm.edu/10.1037/0003-066X.55.1.68

Sadler, J. (2003). Effectiveness of student admission essays in identifying attrition. *Nurse Education Today*, *23*(8), 620–627. https://doi.org/10.1016/S0260-6917(03)00112-6

Safron, D. J., Schulenberg, J. E., & Bachman, J. G. (2001). Part-Time Work and Hurried Adolescence: The Links among Work Intensity, Social Activities, Health Behaviors, and Substance Use. *Journal of Health and Social Behavior*, *42*(4), 425–449. https://doi.org/10.2307/3090188

Santelices, M. V., & Wilson, M. (2010). Unfair Treatment? The Case of Freedle, the SAT, and the Standardization Approach to Differential Item Functioning. *Harvard Educational Review*, *80*(1), 106–134. https://doi.org/10.17763/haer.80.1.j94675w001329270

Sax, L. J., Astin, A. W., & Avalos, J. (1999). Long-Term Effects of Volunteerism During the Undergraduate Years. *The Review of Higher Education*, *22*(2), 187–202.

Sax, L. J., & Harper, C. E. (2007). Origins of the Gender Gap: Pre-College and College Influences on Differences Between Men and Women. *Research in Higher Education*, *48*(6), 669–694. https://doi.org/10.1007/s11162-006-9046-z

Scherer, D. L. (1985). Measuring the Measurements: A Study of Evaluation of Writing: An Annotated Bibliography.

Schmalz, G. M., Rahr, R. R., & Allen, R. M. (1990). The Use of Pre-Admission Data to Predict Levels of Success in Selected Allied Health Students. *The Occupational Therapy Journal of Research*, *10*(6), 367–376. https://doi.org/10.1177/153944929001000606

Schmidt, F. L., & Hunter, J. (2004). General Mental Ability in the World of Work: Occupational Attainment and Job Performance. *Journal of Personality and Social Psychology*, *86*(1), 162–173. https://doi.org/10.1037/0022-3514.86.1.162

Schmitt, N., Keeney, J., Oswald, F. L., Pleskac, T. J., Billington, A. Q., Sinha, R., & Zorzie, M. (2009). Prediction of 4-year college student performance using cognitive and noncognitive predictors and the impact on demographic status of

admitted students. *Journal of Applied Psychology*, *94*(6), 1479–1497. http://dx.doi.org/10.1037/a0016810

Sedlacek, W. E. (2003). Alternative Admissions and Scholarship Selection Measures in Higher Education. *Measurement & Evaluation in Counseling & Development*, *35*(4), 263.

Shaw, E. J., Mattern, K. D., & Patterson, B. F. (2011). Discrepant SAT Critical Reading and Writing Scores: Implications for College Performance. *Educational Assessment*, *16*(3), 145–163. https://doi.org/10.1080/10627197.2011.604241

Sheldon, K. M., Ryan, R. M., Rawsthorne, L. J., & Ilardi, B. (1997). Trait Self and True Self: Cross-Role Variation in the Big-Five Personality Traits and Its Relations With Psychological Authenticity and Subjective Well-Being. *Journal of Personality and Social Psychology*, *73*(6), 1380–1393. https://doi.org/10.1037/0022-3514.73.6.1380

Simons, J., Dewitte, S., & Lens, W. (2004). The role of different types of instrumentality in motivation, study strategies, and performance: Know why you learn, so you'll know what you learn! *British Journal of Educational Psychology*, *74*, 343–360.

Siu, E., & Reiter, H. I. (2009). Overview: what's worked and what hasn't as a guide towards predictive admissions tool development. *Advances in Health Sciences Education*, *14*(5), 759. https://doi.org/10.1007/s10459-009-9160-8

Snellman, K., Silva, J. M., Frederick, C. B., & Putnam, R. D. (2015). The Engagement Gap: Social Mobility and Extracurricular Participation among American Youth. *The ANNALS of the American Academy of Political and Social Science*, *657*(1), 194–207. https://doi.org/10.1177/0002716214548398

Sternberg, R. J. (2004). WICS: A Model of Educational Leadership. *Educational Forum, The*, *68*(2), 108–114.

Sternberg, R. J. (2006). The Rainbow Project: Enhancing the SAT through assessments of analytical, practical, and creative skills. *Intelligence*, *34*(4), 321–350. https://doi.org/10.1016/j.intell.2006.01.002

Sternberg, R. J. (2009). The Rainbow and Kaleidoscope Projects. *European Psychologist*, *14*(4), 279–287. https://doi.org/10.1027/1016-9040.14.4.279

Sternberg, R. J. (2010). Assessment of gifted students for identification purposes: New techniques for a new millennium. *Learning and Individual Differences*, *20*(4), 327–336. https://doi.org/10.1016/j.lindif.2009.08.003

Sternberg, R. J., & Lubart, T. I. (1996). Investing in creativity. *American Psychologist, 51*(7), 677–688. http://dx.doi.org.ezproxy.uvm.edu/10.1037/0003-066X.51.7.677

Strayhorn, T. L. (2014). What Role Does Grit Play in the Academic Success of Black Male Collegians at Predominantly White Institutions? *Journal of African American Studies, 18*(1), 1–10. https://doi.org/10.1007/s12111-012-9243-0

Stukas, A. A., Snyder, M., & Clary, E. G. (1999). The Effects of "Mandatory Volunteerism" on Intentions to Volunteer. *Psychological Science, 10*(1), 59–64. https://doi.org/10.1111/1467-9280.00107

Sullivan, W., & Rosin, M. S. (2008). *A new agenda for higher education: shaping a life of the mind for practice / William M. Sullivan and Matthew S. Rosin ; foreword by Lee S. Shulman and Gary D. Fenstermacher.* (1st ed..). San Francisco: Jossey-Bass.

Thomas, L. L., Kuncel, N. R., & Credé, M. (2007). Noncognitive Variables in College Admissions: The Case of the Non-Cognitive Questionnaire. *Educational and Psychological Measurement, 67*(4), 635–657. https://doi.org/10.1177/0013164406292074

Tracey, T. J., & Sedlacek, W. E. (1984). Noncognitive Variables in Predicting Academic Success by Race. *Measurement and Evaluation in Guidance, 16*(4), 171–178. https://doi.org/10.1080/00256307.1984.12022352

Trevino, N. N., & DeFreitas, S. C. (2014). The relationship between intrinsic motivation and academic achievement for first generation Latino college students. *Social Psychology of Education, 17*(2), 293–306. https://doi.org/10.1007/s11218-013-9245-3

Vallerand, R., & BIssonnette, R. (1992). Intrinsic, Extrinsic, and Amotivational Styles as Predictors of Behavior: A Prospective Study. *Journal of Personality, 60*(3), 599–620. https://doi.org/10.1111/j.1467-6494.1992.tb00922.x

Vallerand, R. J. (1989). Vers une méthodologie de validation trans-culturelle de questionnaires psychologiques: Implications pour la recherche en langue française. *Canadian Psychology/Psychologie Canadienne, 30*(4), 662.

Vallerand, Robert J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., & Vallieres, E. F. (1992). The Academic Motivation Scale: A Measure of Intrinsic, Extrinsic, and Amotivation in Education. *Educational and Psychological Measurement, 52*(4), 1003–1017. https://doi.org/10.1177/0013164492052004025

Van Willigen, M. (2000). Differential Benefits of Volunteering Across the Life Course. *The Journals of Gerontology: Series B*, *55*(5), S308–S318. https://doi.org/10.1093/geronb/55.5.S308

Veal, L. R., & Hudson, S. A. (1983). Direct and Indirect Measures for Large-Scale Evaluation of Writing. *Research in the Teaching of English*, *17*(3), 290–296.

Vught, F. A. V., & Westerheijden, D. F. (1994). Towards a general model of quality assessment in higher education. *Higher Education*, *28*(3), 355–371. https://doi.org/10.1007/BF01383722

Wallach, M. A., & Wing, C. W. (1969). The Talented Student; A Validation of the Creativity-Intelligence Distinction.

Weller, N. F., Kelder, S. H., Cooper, S. P., Basen-Engquist, K., & Tortolero, S. R. (2003). School-Year Employment Among High School Students: Effects on Academic, Social, and Physical Functioning. *Adolescence; Roslyn Heights*, *38*(151), 441–458.

Whitt, E. J., Edison, M. I., Pascarella, E. T., Terenzini, P. T., & Nora, A. (2001). Influences on Students' Openness to Diversity and Challenge in the Second and Third Years of College. *The Journal of Higher Education*, *72*(2), 172–204. https://doi.org/10.1080/00221546.2001.11778877

Wilson, J., & Musick, M. (1997). Who Cares? Toward an Integrated Theory of Volunteer Work. *American Sociological Review*, *62*(5), 694–713. https://doi.org/10.2307/2657355

Zwick, R. (2013, April 15). Is the SAT a "Wealth Test"? The Link between Educational Achievement and Socioeconomic Status. https://doi.org/10.4324/9780203463932-24

Appendices

Appendix A
Preliminary Essay Assessment Rubric

	Definition	Subconstructs	Subconstruct Definition	Evidence			
				No Evidence Provided	Low (1,2)	Mid (3, 4, 5)	High (6, 7)
Grit	Perseverance and passion for long-term goals. Working strenuously towards challenges, maintaining effort and interest over years despite failure, adversity, and plateaus in progress (Duckworth et al., 2007)	Consistency of Interest (Duckworth et al., 2007)	Sustained interest and focus on projects over time. Stamina with which one pursues goals.	no evidence provided	Student becomes interested in new pursuits every semester; Student may set a goal, but frequently pursues different goals; Student has difficulty maintaining focus on projects that take more than a few months to complete	Student has dislayed some consistency in their pursuits, but may not have a single interest they have followed throughout high school; Student displays ability to set goals and sometimes pursues them to completion; Student displays a high level of interest in an idea or project, but may not have built upon this interest over the course of	Student has clearly identified interests and pursued them throughout their high school years; Student displays ability to actively set goals and pursue them with focus and; Student displays a strong ability to focus on projects that span mutltiple semesters or years
		Perseverance of effort (Duckworth et al., 2007)	Enduring effort with zeal for what one is pursuing regardless of immediate feedback. Persuit of a superordinate goal on a longer, more abstract time-scale, despite setbacks (Duckworth et al., 2007)	no evidence provided	Student cannot sustain effort to complete goals; Student displays little followthrough on projects or activities and seems; Student does not seem to embrace hard work or challenges	Student sets goals, but struggles to sustain effort when faced with setbacks; student displays followthrough on some goals and may be moderately discouraged by; student seems capable of hard work, but may be intimidated by challenges	Student may have achieved a goal that took years of work despite multiple perceived or reported setbacks; student does not seem discouraged by failure and has experienced adversity in; Student presents as a hard worker and somone who does not shy away from challenges
Intrinsic Motivation	Doing an activity for itself and the pleasure and satisfaction derived from participation (Deci & Ryan, 1985)	Preference for challenging school work (Leper et al., 2005)	Student's preference for school work that is challenging versus assignments that can be accomplished	no evidence provided	Student does not dipslay a desire to challenge themselves academically	Student occassionallly challenges themselves academically	Student constantly challenges themselves academically
		Engagement, exploration, curiosity (Leper et al., 2005; Vallerand et al., 1992)	The extent to which motivation is measured by engagement, a desire to explore, and personal curiosity	no evidene provided	Student does not display a curiosity to learn or explore new academic concepts	Student displays moderate curiosity to explore new academic concepts	Student shows high motivation to engage with new academic concepts and shows or articulate academic curiosity
		Independent Mastery (Leper et al., 2005)	Ability and desire to master material independently without relying on instruction	no evidence provided	Student relies heavily on structured environments to learn academic concepts	Student displays some self-guided learning strategies through extracurriculars or academic projects	Student shows stong ability to learn new academic concepts without instruction and seeks out self-learning opportunities through multiple avenues
		Learning orientation vs performance orientation. Intrinsic	Motivation through desire to learn about a certain topic versus motivation to demonstrate one's	no evidence provided	Student appears to be motivated mostly by external performance (e.g. grades, awards, expectations)	Student displays motivation to learn, but seems to repspond more strongly to external accomplishments (e.g. chose advanced course work	Student displays high motivation to learn and motivation persists regardless of external feedback or awards
Creativity	Accomplishments or creation of products that are viewed as useful, valued, original, or novel (Dollinger, 2011)	Creativity	Ability to see problems in new ways and escape the bounds of conventional thinking	no evidence provided	Student displays conventional approaches to tasks and problem solving	Student displays mild ability to see problems in new ways. Makes attempts to approach concepts in unconventional	Student displays strong ability to approach problems in new ways and see outside conventional thinking
		Originality	Cleverness, humor, originality (Latent creativity variables from Sternberg, 2006)	no evidence provided	Student does not demonstrated originality in their work or writing	Student demonstrates some originality in their work or writing. Some evidence can be found through the presence of	Student demonstrates a great deal of originality in their academic pursuits and or writing. Student aptly
Leadership	Demonstrating skills in a group, such as motivating others, coordinating groups and tasks, serving as a representative for the group, or otherwise performing a managing role in a group. (Schmidtt et al., 2009; Bruggink & Gambhir, 1996 - provide rating scale from Willingham and Breeland 1982 book Personal Qualities and College Admissions)	Leadership: Community, School, or Athletic	Extracurriculars outside of school sponsored activities (non-athletic)	no evidence provided	Some involvement but may not extend beyond simple participation	Special local achievement or membership in a prominent	Exceptional individual achievement, more than local recog
Community Engagement	Demonstrates interest and commitment to community activities beyond surface level or self-serving activities	N/A	N/A	no evidence provided	Student displays some engagemnt in the greater community outside of their high school. Participation is not self-serving	student displays a good deal of engagement in their greater community beyond the local level. Participation is not self-serving.	Student displays a high level of engagement in their community and has had national or international engagement. Participation is not self-serving
Cultural Fluency	a person who is able to move outside of their natural comfortable community.	N/A	N/A	no evidence provided	Participation in some type of activity outside of their natural born community. Cross cultural bridgebuilding.		

81

Field Testing Interview Protocol

Date: _____

Purpose:
The purpose of this interview is to obtain your feedback on an early draft of a rubric that may be used to help admissions officers at UVM operationalize their essay evaluation process.

Permission to abstain:
If at any point you wish to abstain from answering a question, please let me know. If you wish to terminate the interview at any point, you may do so.

Retention of feedback:
I will be taking notes for the purposes of altering and improving my matrix. I will then destroy the notes after integrating the feedback.

Anonymity:
Your feedback is being used to purely to improve an evaluative matrix and the only reference to our interview will be "I spoke with five admissions professionals at peer or peer aspirant institutions and obtained their feedback on an early draft of my rubric. All of these individuals spoke with me as an extension of their professional capacity." There will be no personally identifying information.

If you do not have any questions, may we begin?

Interview Protocol

1. As an experienced admissions officer, what makes a good essay in your mind?
    a. What makes a poor essay?
2. In your professional opinion, what does the essay assist you in measuring?
3. Are there any themes or constructs you actively look for when reading an essay? (for example, creativity, perseverance?)
4. Are there any items you think are easily measured in an essay?
5. Does your office currently use a rubric to evaluate essays?
    a. If yes, are you able to share with me how your rubric was constructed and what it is designed to measure?

At UVM we've identified the following constructs as desirable in an applicant:
1. Grit
2. Intrinsic Motivation
3. Creativity
4. Leadership
5. Community Engagement

I'm going to go through each of these constructs one at a time and ask for your professional interpretation of their definition and how you might identify evidence of each construct.

Grit
1. How would you define Grit in the context of admissions?
2. In your professional opinion, is Grit a valuable trait in an applicant? Why or why not?
3. Do you think it is measurable in an essay? How might it show up?
4. If you had to evaluate a student's level of grit, what would low evidence of grit look like to you? What would high evidence of grit look like to you?

Intrinsic Motivation
1. How would you define Intrinsic Motivation in the context of admissions?
2. In your professional opinion, is intrinsic motivation a valuable trait in an applicant? Why or why not?
3. Do you think it is measurable in an essay? How might it show up?
4. If you had to evaluate a student's level of intrinsic motivation, what would low evidence of intrinsic motivation look like to you? What would high evidence of intrinsic look like to you?
5. Can intrinsic motivation be measured in an essay?

Creativity
1. How would you define creativity in the context of admissions?
2. In your professional opinion, is creativity a valuable trait in an applicant? Why or why not?
3. Do you think it is measurable in an essay? How might it show up?
4. If you had to evaluate a student's level of creativity what would low evidence of creativity look like to you? What would high evidence of creativity look like to you?

Leadership
1. How would you define leadership in the context of admissions?
2. In your professional opinion, is leadership a valuable trait in an applicant? Why or why not?
3. Do you think it is measurable in an essay? How might it show up?
4. If you had to evaluate a student's level of leadership what would low evidence of leadership look like to you? What would high evidence of leadership look like to you?

Community Engagement
1. How would you define community engagement in the context of admissions?
2. In your professional opinion, is community engagement a valuable trait in an applicant? Why or why not?
3. Do you think it is measurable in an essay? How might it show up?
4. If you had to evaluate a student's level of engagement what would low evidence of engagement look like to you? What would high evidence of engagement look like to you?

Cultural Fluency
1. How would you define cultural fluency in the context of admissions?
2. In your professional opinion, is cultural fluency a valuable trait in an applicant? Why or why not?
3. Do you think it is measurable in an essay? How might it show up?
4. If you had to evaluate a student's level of cultural fluency what would low evidence of cultural fluency look like to you? What would high evidence of cultural fluency look like to you?

Of the non-cognitive factors we discussed, which do you think will be most easily measured in an essay? Do you think one factor is more important than another?

We will now take some time to review the preliminary rubric together

Can any of these non-cognitive factors be defined without sub-constructs?
Is one particular sub-construct the best proxy for the overall definition of the characteristic?

Do you have any questions for me?