

THE DEVELOPMENT AND VALIDATION
OF A VIGNETTE-BASED ACADEMIC
GRIT SCALE

by

MITCHELL PORTER

SARA TOMEK, COMMITTEE CHAIR

YOUN-JENG CHOI
CECIL ROBINSON
RANDALL SCHUMACKER
STEFANIE A. WIND

A DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in the
Department of Educational Studies
in Psychology, Research
Methodology, and Counseling
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2019

Copyright Mitchell Porter 2019
ALL RIGHTS RESERVED

ABSTRACT

This study conducted a development and validation of a vignette-based grit scale (Grit-V). The purpose was to determine if the use of vignettes provided more validity evidence than short simple-sentence item endorsements. The data were gathered from a sample population of undergraduate students from The University of Alabama. Focus group data were gathered to help provide insight during the item development process, and an expert panel was used to help select which items were most appropriate based on the item content. A pilot study was conducted to see how the initial item pool was functioning. The results from the pilot facilitated in the selection of the final items to form the Grit-V. The study therefore investigated whether the Grit-V provided more validity evidence than the existing measure of grit (Grit-S). The study further investigated the relationships between grit and other non-cognitive constructs (student engagement, academic motivation, and the Big Five personality dimensions). Additionally, the study determined the predictive strength of grit and the other variables on student success, which was measured by student GPA, classes dropped, and changed majors. Finally, the study investigated whether socioeconomic status was a significant moderator variable between grit and GPA. The results showed that the Grit-V provided more evidence of validity than the Grit-S. Grit and conscientiousness were the strongest predictors of GPA, and grit, academic motivation, and student engagement were the strongest predictors of dropping classes. Grit was not a significant predictor for changed majors. Additionally, SES was not a significant moderator variable in the predictive model.

Keywords: grit, vignettes, scale development, non-cognitive skills, assessment

DEDICATION

I dedicate the completion of my dissertation to my amazing family for supporting me throughout this entire process. My parents, Robert and Joan Porter, have always believed in me and provided me with the courage I needed to finally accomplish my goals. My beloved wife, Dr. Sijia Zhang, stood by my side when I needed her the most. My amazing son, Aiden, gave me strength and motivation to complete my doctoral program. Without the love and support of my family, this journey would not have been possible.

ACKNOWLEDGMENTS

I have been very fortunate to have such a thoughtful dissertation committee. I would like to thank Drs. Sara Tomek, Stefanie Wind, Cecil Robinson, Randall Schumacker, and Youn-Jeng Choi for their support not only with my dissertation, but also throughout my entire doctoral program. Without them, I would not have been able to push through the difficult times and finish. This has been an incredible experience, and their support will always be remembered.

CONTENTS

ABSTRACT	ii
DEDICATION	iii
ACKNOWLEDGMENTS	iv
LIST OF TABLES.....	xii
LIST OF FIGURES.....	xiv
CHAPTER I: INTRODUCTION	1
The Quandary of Defining Academic Success.....	1
Cognitive and Non-cognitive Skills and Their Role in Success.....	2
Why Grit?	3
Improving Self-assessment with Vignette-based Measurement.....	4
Confirming the Factor Structure of Grit as a Construct.....	5
Socioeconomic Status as a Moderator Variable.....	6
Definitions of Concepts.....	7
Research Questions	8
CHAPTER II: REVIEW OF THE LITERATURE	9
How Academic Success Has Been Measured and Defined in Recent Literature.....	11
Cognitive and Non-cognitive Skills.....	15
Conscientiousness and Goal Orientation.....	20
Grit	21

Grit, Race, and SES.....	23
Exploring the Role of SES in Educational Assessments.....	23
Concerns with Traditional Self-report Assessments.....	24
Vignettes over Traditional Item Endorsement in Measurement.....	26
CHAPTER III: METHODS.....	30
Step 1: Focus Groups	31
Participants.....	31
Data Collection	32
Instrument	33
Procedure	33
Step 2: Item and Scale Development	34
Step 3: Pilot Study and Preliminary Analysis	35
Procedure	35
Participants.....	36
Data Analysis	37
Partial credit Rasch model analysis	37
Confirmatory and exploratory factor analysis.....	37
Validity.....	38
Revisions and final scale form.....	39
Step 4: Final Item Pool and Analyses	39
Participants.....	39
Measures and Instruments.....	40
Grit	40

Motivation	41
Student engagement	41
Academic success	41
Personality dimensions.....	42
Socioeconomic status.....	42
Procedure	42
Analysis	43
Research Question 1	43
Research Question 2	45
Research Question 3	48
Research Question 4	50
CHAPTER IV: RESULTS.....	51
Item and Scale Development.....	51
Focus Groups	51
Content Validity	56
Pilot Study Results.....	57
Partial Credit Rasch Model.....	57
Construct Validity	62
Convergent Validity	64
Final Item Revisions and Selection.....	66
Final Study.....	67
Descriptive Statistics	67
Research Question 1.....	69

Reliability.....	69
Partial Credit Rasch Model Analysis.....	69
Categorical Ordering and Precision.....	75
Dimensionality	78
ANOVA	82
Completion Time Differences.....	83
Research Question 2.....	83
Cronbach’s Alphas	83
Comparison of Grit-V and Grit-S on GPA	83
Comparison of Grit-V and Grit-S on Dropped Classes	84
Comparison of Grit-V and Grit-S on Changed Majors	85
Comparison between Grit-V, Grit-S, NSSE, AMS-C, BFI Personality Dimensions on GPA	86
Comparison between Grit-V, Grit-S, NSSE, AMC-C, BFI Personality Dimensions and Classes Dropped	90
Comparison between Grit-V, Grit-S, NSSE, AMC-C, BFI Personality Dimensions and Changed Majors.....	92
MANOVA.....	94
Summary of RQ2 Results	96
Research Question 3.....	97
Criterion Validity	97
Construct Validity	98
Convergent Validity	98
Content Validity	101
Summary of RQ3 Results	101

Research Question 4.....	101
CHAPTER V: DISCUSSION.....	104
Item Development Process.....	104
Focus Groups.....	104
Psychometric Properties.....	105
Reliability.....	105
Partial credit Rasch model.....	105
<i>Model data fit</i>	105
<i>Facets</i>	106
<i>DIF</i>	107
<i>Category ordering and precision</i>	107
Validity.....	108
Criterion validity.....	108
Construct validity.....	109
Convergent validity.....	109
Content validity.....	110
Completion Times for the Grit-V and Grit-S.....	110
The Predictive Ability of Grit and the Other Constructs on Academic Success.....	111
The Relationships Between Grit and the Other Constructs.....	113
Grit and Student Engagement.....	113
Grit and Academic Motivation.....	114
Grit and Conscientiousness.....	114
Grit and Neuroticism.....	115

Grit and Extraversion.....	116
Grit and Agreeableness.....	116
Grit and Openness to Experience.....	116
Grit-V Guide.....	117
Potential Admissions Criteria.....	117
Counseling.....	118
Future Grit Research.....	118
Interpretations.....	119
Low.....	119
Medium.....	120
High.....	120
Very high.....	120
Making admission decisions.....	121
Limitations and Future Directions.....	122
Grit and Degree Completion.....	122
Grit and Other Student Populations.....	122
Longitudinal Grit Research.....	123
Grit and Self-regulation.....	123
Gender Differences.....	124
Grit and SES.....	124
Grit and Children.....	124
Socially Desirable Responses.....	125
The Use of Grit Measurement in Business.....	125

The Use of Vignettes in Measurement	126
Final Remarks	127
REFERENCES	128
APPENDIX A: FOCUS GROUP GUIDE	140
APPENDIX B: GRIT-V SCALE (INITIAL ITEM POOL FOR PILOT STUDY)	143
APPENDIX C: GRIT-V SCALE (FINAL FORM).....	150
APPENDIX D: GRIT-S SCALE.....	156
APPENDIX E: ACADEMIC MOTIVATION SCALE (AMS-C 28).....	158
APPENDIX F: NATIONAL SURVEY OF STUDENT ENGAGEMENT (NSSE).....	161
APPENDIX G: ACADEMIC SUCCESS QUESTIONNAIRE	163
APPENDIX H: BIG FIVE INVENTORY (BFI)	165
APPENDIX I: IRB APPROVAL	167

LIST OF TABLES

2.1.	Summary of Cognitive Tests	10
2.2.	Summary of Non-cognitive Assessments.....	11
3.1.	Demographic Information for Focus Groups	32
3.2.	Demographic Information for Pilot Study.....	36
3.3	Demographic Information for Final Analysis	40
3.4.	Summary of Regression Models.....	47
3.5.	Model Summary for Moderator Analysis	50
4.1.	Summary of Expert Panel Review of Vignettes.....	57
4.2.	Summary of Calibrations and Model-data Fit Statistics	58
4.3.	Summary of Item Estimates	60
4.4.	Summary of Student Estimates.....	61
4.5.	Summary of Gender Differences	62
4.6.	Item Loadings for the Two-Factor Model.....	63
4.7.	Principle Component Analysis Factor Solution	64
4.8.	Summary of Correlation Results with Grit-V.....	65
4.9.	Descriptive Statistics.....	68
4.10	Correlation Matrix.....	69
4.11.	Summary of Calibrations and Model-data Fit Statistics	71
4.12.	Summary of Item Estimates	72
4.13.	Summary of Gender Differences	74

4.14.	Threshold Distances for Grit-V Items.....	76
4.15.	Frequencies of Category Observations	77
4.16.	Average Logit-scale Locations Across Items and Categories.....	78
4.17.	Grit-V Item Loadings for the Two-factor Model.....	79
4.18.	Grit-S Item Loadings for the Two-factor Model	80
4.19.	Summary of the Single-factor CFA Model for Grit-V and Grit-S	81
4.20.	Score Ranges for the Grit-V	82
4.21.	Summary of Grit-Level Estimates	83
4.22.	Summary Results of the Comparison of Grit-V and Grit-S on GPA.....	84
4.23.	Summary Results of the Comparison between Grit-V and Grit-S on Dropped Classes...	85
4.24.	Summary Results of the Comparison between Grit-V and Grit-S on Changed Majors ...	86
4.25.	Summary Results of the Comparison between Grit-V, Grit-S, NSSE, AMS-C, BFI Personality Dimensions on GPA	89
4.26.	Summary Results of the Comparison between Grit-V, Grit-S, NSSE, AMC-C, BFI Personality Dimensions and Classes Dropped	91
4.27.	Summary Results of the Comparison between Grit-V, Grit-S, NSSE, AMC-C, BFI Personality Dimensions and Changed Majors	93
4.28.	Summary of MANOVA Results.....	96
4.29.	Summary of Correlation Results with Grit-V.....	99
4.30.	Summary of Correlation Results with Grit-S	100
4.31.	Summary Results of Moderator Analysis	102
4.32.	Summary of Education Level Estimates for Grit.....	103
5.1.	Point Breakdown for the Grit-V	119

LIST OF FIGURES

4.1. Graphical representation of item 1 residuals	75
---	----

CHAPTER I: INTRODUCTION

Measuring academic success is an issue in education that has been exhausted, yet researchers are still inconclusive in identifying what constructs are the most telling of an individual's potential academic success. More specifically, what the factors are that determine whether a student is, or will be, successful, is a widely-debated topic. Researchers have proposed several different ways to measure and predict academic success. Unfortunately, the majority of the proposed methods of assessment fall under self-report measures that rely solely on item endorsement. I believe that by examining grit, a component of academic success through the use of a vignette-based measurement tool, we can get a more balanced predictor of a person's academic success. The study examined the development and validation of a vignette-based measurement tool that assesses an individual's level of grit to predict their academic success.

The Quandary of Defining Academic Success

Indeed, defining academic success is a difficult task that has been long attempted by researchers yet never completely agreed upon. Tinto and Pusser (2006) have provided an open interpretation for academic success: "Our discussion leaves open the definition of success other than to imply that without learning there is no success and, at a minimum, success implies successful learning in the classroom" (p. 8). While philosophical in nature, I targeted a more concrete meaning of student success. This will be expanded upon in Chapter II.

In this study, I do not take on the task of defining academic success in its entirety. Rather, I developed and validated a new vignette-based scale (Grit-V) to measure grit as a component of

academic success. Grit is a non-cognitive skill, defined as trait level perseverance and conscientiousness (Duckworth, Peterson, Matthews, & Kelly, 2007). In most cases, these contexts around grit are related to classroom performance (e.g., GPA, dropout rates, degree completion) and success in academia (e.g., securing a faculty position in a higher education institution) as well as life achievements. While grit can be used in many contexts, this research focused on grit in academic settings.

Previous research has shown positive relationships between non-cognitive skills such as grit and academic success. Hochanadel and Finamore (2015) have discussed the positive implications of teaching non-cognitive skills, such as growth mindset and grit, in the classroom. They found that when students are taught what grit is and how it can benefit their lives, they are more likely to persist in the face of adversity. The advantages of the teaching of grit is not limited to students; faculty and parents would also benefit greatly from having a firm understanding of grit.

Cognitive and Non-cognitive Skills and Their Role in Success

In modern psychology, a popular concept surrounding success is based on the cognitive hypothesis (Tough, 2013), which suggests that the primary factor in determining a person's future success lies within their cognitive abilities. Cognitive skills (or brain functions) are related to how a person learns, remembers, and solves problems. An example would be motor skills, which is the ability to manipulate objects using bodies. One possible reason this particular theory gained so much traction and attention is the perceived ease of measuring a person's intelligence quotient (IQ). In the past, the IQ was used as the primary predictor of a person's life success, and it did so with a respectable amount of accuracy (Gottfredson, 1997; Hartigan & Wigdor, 1989).

However, it is believed that other factors may attribute to a person's success. Terman, Oden, and Bayley (1947) commented in their work on gifted children that additional factors, other than IQ, accounted for children's future success. What Terman et al. noted in their study was that non-cognitive abilities are significant contributing factors in success. Non-cognitive abilities develop over time and include things such as attitude, behavior, and motivation. There has been some resistance to Terman et al.'s asserted importance of non-cognitive abilities. Through this research, I explored the relationship between grit and students' success in their education.

Why Grit?

I chose to explore the relationships between grit and academic success because psychologists and researchers interested in non-cognitive skills have established that grit has positive correlations with a person's long-term success (Duckworth et al., 2007). After viewing Angel Duckworth's TED Talk highlighting her research on grit (Duckworth, 2013), it became clear that grit was an alternative way to measure person's success. This study therefore hypothesizes that a vignette-based scale measuring grit can be more useful than the current measure (Grit-S, Duckworth, & Quinn, 2009) in predicting a student's success in the classroom (short-term success) and persevering through their program (long-term success). However, since this is a cross-sectional study as opposed to a longitudinal study, the study only took an individual's snapshot of grit (i.e., short-term goals).

Grit, as described by Duckworth (2007), "overlaps with achievement aspects of conscientiousness but differs in its emphasis on long-term stamina rather than short-term intensity" (p. 1089). This would lead us to believe that people with higher levels of grit are not only able to finish their current goals, but also to continue to seek additional long-term goals. In

the context of grit in academia, for example, let us consider a student with grit whose goal is to obtain a faculty position at a higher education institution. I believe that this student would be able to accomplish both short-term goals (achieve a high GPA) as well as long-term goals (complete their degree or not changing majors) in the future.

Improving Self-assessment with Vignette-based Measurement

Another goal of this study was to show how some measurement tools can benefit from a vignette-based approach to item writing. Survey-based self-report measurement tools are abundant in the literature (e.g., *Adult Hope Scale* by Snyder et al., 1991; *Grit-S* by Duckworth & Quinn, 2009; *The Self Efficacy Scale* by Sherer, Maddux, Mercandante, Prentice-Dunn, Jacobs, & Rogers, 1982; *The Boredom Proneness Scale* by Farmer & Sundberg, 1986; *The Center for Epidemiologic Studies Depression Scale* by Radloff, 1977). They are cost-effective and efficient to disburse to a large sample. However, the validity of self-report measures has been questioned for years (e.g. Fan, Miller, Park, Winward, Christensen, Grotevant, & Tai, 2006; Austin, Deary, Gibson, McGregor, & Dent, 1998; Fujii, Hennessy, & Mak, 1985). Can we really say that they are providing us the best, most accurate data? Self-reported data will always be a popular choice among social researchers; however, a vignette-based approach to scale development offers a more context-rich pool of items that can better assess the desired construct.

Worthington and Whittaker (2006) conducted a content analysis and provided recommendations for best practices in scale development research. They reveal that much has been researched in terms of statistical analyses being used in psychometric practices (e.g., confirmatory factor analysis, exploratory factor analysis, and structural equation modeling). Other common themes explored include criteria for assessing the factorability, extraction methods, rotation methods, and purposes and criteria for optimizing scale length. The authors did

not however investigate the item development process, such as how the items are worded or created.

While self-assessment has inherent flaws, such as individual response spread in self-report scales (Austin, Deary, Gibson, Murray, McGregor, & Dent, 1998) and self-report bias (Alexander & Becker, 1978), it is typically the most practical approach to data collection in the social sciences. As an example, the current self-assessment survey measure of grit, the Grit-S scale (Duckworth & Quinn, 2009) has even been questioned by its creator as to its validity (Duckworth & Yeager, 2015). Therefore, one of my goals with this research was to improve upon the existing Grit-S scale through the process of developing a vignette-based grit scale (Grit-V) with the hopes of reducing the inherent reference bias found in self-report measurement tools.

Confirming the Factor Structure of Grit as a Construct

The current literature operationalizes grit as a higher-order construct with two lower order facets: perseverance of effort and consistency of interest (Duckworth et al., 2007). These facets refer to a person's tendency to work hard even in the face of adversity (perseverance of effort) and the tendency not to change goals and interests frequently (consistency of interest). Why are these important facets? According to Ericsson, Krampe, and Tesch-Römer (1993), persistence and consistency are both necessary for success. Persistence is necessary because the process of mastering something involves obstacles that the person must persist through. Consistency is important because to achieve mastery of something typically requires many hours of practice. Therefore, it can also be deduced that a person who does not engage in something when facing setbacks and fails to put in the necessary hours is unlikely to achieve mastery or high levels of performance. The Grit-S scale (Duckworth & Quinn, 2009) contains these two facets in the subscales, although the overall grit score is what is primarily used in the literature.

So why is it that researchers have ignored the subscores and focused solely on the overall score? There appears to be two explanations, and both are provided by Duckworth. The first explanation is that the factors by themselves are not as predictive as the summative value of both. Also, neither factor was more predictive of success over the other. The second explanation is based on confirmatory factor analysis. Duckworth and Quinn (2009) claimed that the higher-order structure exhibited significantly better fit indices than a signal factor model. There are statistical issues with these claims, and they will be explored more closely in Chapter II.

Socioeconomic Status as a Moderator Variable

The vast majority of literature on grit and its relation to success has focused on a Western population. However, very few researchers have investigated the effect of grit with students from different cultural backgrounds, nor have relationships between socioeconomic status (SES), grit, and success been explored in any depth. The one of interest in this research is the latter.

Datu, Yuen, and Chen (2017) examined the effect of personality traits, including determination and perseverance, in Asian contexts. They noted that certain items did not load on the higher-order grit construct, and some items did not significantly load to the “Consistency of Interests” dimension. This may be due to the notion of grit having a different cultural meaning in a collectivist society than in Western society. Herold (2015) has also posited that grit is culture-biased and is racially constructed. This research aimed to expand upon previous literature, or lack thereof, and explore SES as a moderator variable for grit as it relates to academic success.

In regard to SES, Duckworth’s recent book *Grit: The power of passion and perseverance* (2016) discussed how people from lower SES groups typically have higher levels of grit. This is because they encounter more hardships than say a more privileged individual from a higher SES, and these hardships are important for grit to help shape one’s personality. However, there is no

statistical evidence to back this up; much of the discussion stems from anecdotal evidence. In this study, SES is included as an important factor of one's level of "grittiness."

Definitions of Concepts

Grit: As defined by Duckworth et al. (2007), grit is trait-level perseverance and passion for long term goals. In the context of this study, grit will retain this meaning; however, it will be used strictly in the context of academia.

Academic Motivation: For the purposes of this study, motivation will be comprised of three concepts: intrinsic motivation, extrinsic motivation, and amotivation. Intrinsic motivation is doing something for itself (the act of doing something provides satisfaction) (Deci and Ryan, 1985). Contrarily, extrinsic motivation is the act of doing something as a means to an end Deci, 1975). Amotivation is the absence of intrinsic and extrinsic motivation.

Student Engagement: Student engagement has been conceptualized by Marks (2000) as "a psychological process, specifically, the attention, interest, investment, and effort students expend in the work of learning" (pp. 154-155). This can be extended to include students' involvement with school. In other words, it is the level of investment and effort put forth to learn in school settings.

Vignette: A vignette is a short story or description of a specific scenario (Alexander and Becker, 1978). This can be in the form of writing, audio, and visual. In the field of psychometrics, and in the context of this paper, each scenario is an item on a scale that portrays a selected element of reality (in this case, grit) and asks participants to respond by selecting an outcome that best fits how they would react to the given scenario (Hughes, 2012).

Socioeconomic status: Socioeconomic status (SES) is a compilation of educational level, annual income, and occupational status. For this study, SES will be assessed by measuring the participants' parents' SES.

Research Questions

The following four research questions were addressed. They included

1. What are the psychometric properties of the Grit-V scale;
2. Does the Grit-V scale better relate to success in academia than the Grit-S scale and Big Five dimensions;
3. Does the use of vignettes provide stronger evidence of validity when measuring grit compared to short sentence items; and
4. Is SES a moderator variable when determining a person's success in relation to their level of grit?

CHAPTER II:

REVIEW OF THE LITERATURE

A long-standing debate in education has been how to determine academic success in students. York, Gibson, and Rankin (2015) provided an analytic review of literature surrounding the notion of academic success, and how researchers and educators have defined it. By and large, the majority of the assessments designed to measure academic success focus on critical thinking skills in one manner or another. The following section outlines the current trends in assessing and defining academic success. Table 2.1 contains a summary table of cognitive assessments and Table 2.2 contains a summary table of non-cognitive assessments; both tables include title, authors, and purpose.

Table 2.1

Summary of Cognitive Tests

Assessment	Authors	Purpose
ETS Proficiency Profile	The Educational Testing Service	To assess general academic skills
Collegiate Assessment of Academic Proficiency- Critical Thinking	American College Testing (1989)	To measure students' skills in clarifying, analyzing, evaluating and extending arguments
California Critical Thinking Dispositions Inventory	Insight Assessment (1992)	To survey the dispositional aspects of critical thinking
California Critical Thinking Skills Test		Measure critical thinking skills
Cornell Critical Thinking Test	Ennis and Millman (1985)	To measure induction, credibility, prediction and experimental planning, deduction, definition, and assumption identification
Problem Solving Inventory	Heepner and Peterson (1982)	To measures decision making in problem solving
The Watson-Glaser Critical Thinking Appraisal	Watson and Glaser (1980)	To measure critical thinking skills
The Critical Thinking Assessment Test		To assess critical thinking and problem solving skills
Emotional Quotient Inventory	Bar-On (2004)	To assess a person's emotional intelligence
Non-cognitive Questionnaire-Revised	Tracey and Sedlacek (1989)	To assess non-cognitive factors of intelligence
Adult Hope Scale	Snyder et al. (1991)	To assess hope in terms of pathway and agency thinking
Cognitive Assessment System	Naglieri and Das (1997)	To evaluate cognitive processing in children
The Montreal Cognitive Assessment	Nasreddine et at. (2005)	To evaluate if a patient exhibits signs of mild cognitive impairment
The Graduate Record Examination	The Educational Testing Service	To evaluate verbal reasoning, quantitative reasoning, critical thinking, and analytic writing skills
The Law School Admission Test	Law School Admission Council	To measure acquired reading and verbal reasoning skills
The Collegiate Assessment of Academic Proficiency	Association of American Medical Colleges	To assess examinees' knowledge of science concepts and principles

Table 2.2

Summary of Non-cognitive Assessments

Assessment	Authors	Purpose
The Multidimensional Scales of Perceived Self-Efficacy	Bandura (1989)	To measure self-efficacy
College Self-Efficacy Inventory	Solberg et al. (1993)	To assess students' perceived level of self-efficacy
Academic Self-Efficacy	Locke and Wood (1988)	To assess students' perceived level of self-efficacy
Self-Regulated Learning Scale	Bandura (1998)	To measure students' ability to engage in self-regulated learning
Student Readiness Inventory	American College Testing	To assess psychological factors that are associated with academic success and college student retention

How Academic Success Has Been Measured and Defined in Recent Literature

The Educational Testing Service (ETS) developed the ETS Proficiency Profile (formerly the Measure of Academic Proficiency Progress, or MAPP), which assesses general academic skills, including critical thinking, reading, writing and mathematics. It is typically administered to undergraduate college students. This test can be used to both assess students' performance as well as the institutional program performance for accreditation. ACT developed the Collegiate Assessment of Academic Proficiency- Critical Thinking (CAAP-CT) in 1989 to measure students' skills in clarifying, analyzing, evaluating and extending arguments. The grounding principle is a student's argumentation skills are indicative of the critical thinking. Insight Assessment developed the California Critical Thinking Dispositions Inventory (CCTDI) in 1992, which includes several subscales, including the Truth-seeking Scale, Open-Mindedness Scale, Analyticity Scale, Systematicity Scale, Critical Thinking Self-Confidence Scale, Inquisitiveness Scale, Maturity of Judgement Scale, and the CCTDI Score Scale. This has been used to survey

the dispositional aspects of critical thinking. Additionally, the California Critical Thinking Skills Test (CCTST) was developed to provide a measure of critical thinking skills, which includes the following scales: Total Score, Analysis, Inference, Evaluation, Deduction, Induction, Interpretation, and Explanation. This assessment seeks to engage the examinees' reasoning skills in a number of different ways, including drawing accurate and warranted inferences, evaluating inferences, and explaining why an inference is either strong or weak. Ennis and Millman (1985) developed the Cornell Critical Thinking Test (CCTT-Z) to measure induction, credibility, prediction and experimental planning, deduction, definition, and assumption identification. The test is comprised of two levels: Level Z is geared toward college students, and X is aimed at students in grades 4 – 14. Heppner and Peterson (1982) developed the Problem Solving Inventory (PSI) which included three subscales: Problem-solving Confidence, Approach Avoidance Style, and Personal Control. This assessment was designed to evaluate the examinees' ability to make decisions related to problem solving. The Watson-Glaser Critical Thinking Appraisal (Watson & Glaser, 1980) is an assessment designed to measure critical thinking skills. For this scale, examinees are provided reading passages, and asked to evaluate them in terms of their problems, statements, arguments, and interpretations. Gadzella, Hogan, Masten, Stacks, Stephens and Zascavage (2006) developed a short form of the scale; the original consisted of 80 items to be completed in 60 minutes, while the short form had 40 items and is to be completed in 45 minutes. The Critical Thinking Assessment Test (CAT) was developed to assess and promote the improvement of critical thinking and dealt with real-world application of problem-solving skills. Rather than multiple choice or Likert scale items, the questions were short answer essay style.

While all the previous examples are widely used and incorporate an individual's critical thinking skills, not all assessments developed to measure academic success are designed in this way. For example, Bar-On (2004) created the Emotional Quotient Inventory (EQI). The EQI is a self-reported measure developed to assess a person's emotional intelligence. In addition to an overall emotional quotient (EQ) it provides scores for five composite scales and 15 subscales (Bar-On, 2006). Further scales (EQ-I, EQ-360, and EQ-i:YV) were developed to assess the Bar-On model of emotional-social intelligence. Tracey and Sedlacek (1989) developed the non-cognitive Questionnaire-Revised (NCQ-R); the factor structure was revised across samples of black and white college students. Snyder et al. (1991) created the 12-item adult hope scale based on Likert scale categories. Four items measured pathways thinking, four measured agency thinking, and four were fillers. The scale is meant to measure Snyder's cognitive model of hope, defined as "a positive motivational state that is based on an interactively derived sense of successful (a) agency (goal-directed energy), and (b) pathways (planning to meet goals)" (Snyder, Irving & Anderson, 1991, p. 287). Clearly, many researchers see the need to measure something beyond academic intelligence.

In terms of cognitive assessment, much of the scale development in this area focuses on identifying learning impairment, as well as social and behavioral disorders. The Cognitive Assessment System (Naglieri & Das, 1997), or CAS, is designed to evaluate cognitive processing in children ages five to seventeen. The scale is derived from the Planning, Attention, Simultaneous, and Successive (PASS) theory, and the four processing areas comprise the four scales that make up the CAS. The Montreal Cognitive Assessment (MoCA) was designed as a brief screening tool to evaluate if a patient exhibits signs of mild cognitive impairment (Nasreddine et al., 2005). A number of assessments have been designed to diagnose obsessive

compulsive disorder (e.g., Brown et al., 1995; Freeston et al., 1993, 1995; Hoekstra, 1995; Sookman & Pinard, 1995; Steketee et al., 1996). Regarding measurement, these scales are all self-reported multiple choice item endorsement.

Other approaches to academic achievement focus on the attainment of learning outcomes. ETS has continually been developing the Graduate Record Examination (GRE) as a way to evaluate potential graduate students' verbal reasoning, quantitative reasoning, critical thinking, and analytic writing skills. This is often used by institutions as one of the factors when making admission decisions. The Law School Admission Test (LSAT), developed by the Law School Admission Council, provides a standard measure of acquired reading and verbal reasoning skills that law schools can use as a factor in deciding whether a candidate is suitable for a particular institute's program. The Collegiate Assessment of Academic Proficiency (CAAP) was created as a standardized nationally normed assessment program from ACT, designed to enable higher education institutions to assess, evaluate, and enhance student learning outcomes, as well as general education learning outcomes. In the field of medicine, the Association of American Medical Colleges developed the Medical College Admission Test (MCAT) as a standardized exam that assesses the examinees' knowledge of science concepts and principles prior to enrolling in medical school. The scores are reported in verbal reasoning, physical and biological sciences, as well as a writing sample. While writing is a significant component of the aforementioned scales, the core of the exams are multiple choice item format.

A number of scales take a different approach, measuring factors related to student success rather than student success itself. Bandura (1989) was the first to develop a scale meant to capture students' self-efficacy. The Multidimensional Scales of Perceived Self-Efficacy (MSPSE) is comprised of 57 items with nine subscales. To expand upon Bandura's work,

Solberg et al. (1993) developed the College Self-Efficacy Inventory (CSEI), again, with the goal of capturing students' perceived level of self-efficacy. Similarly, another self-efficacy scale (Academic Self-Efficacy) was created by Locke and Wood (1988), with the same goal of measuring students' perceived level of self-efficacy. As a further expansion on his previous work, Bandura (1998) designed the Self-Regulated Learning Scale (SRL) in an attempt to measure students' ability to engage in self-regulated learning. This is part of the MSPSE. Lastly, the Student Readiness Inventory (SRI) examines the psychosocial factors that are associated with academic success and college student retention. The assessment consists of ten scales including academic discipline, academic self-confidence, commitment to college, communication skills, steadiness, general determination, goal striving, social activity, social connection, and study skills. Again, the overwhelming majority of these assessments require participants to endorse multiple choice items in order to assess their non-cognitive abilities.

Cognitive and Non-cognitive Skills

When considering academic success, it is often argued that students need to possess both cognitive and non-cognitive skills (Komarraju, Ramsey, & Rinella, 2013). Cognitive skills are typically considered innate abilities, such as reasoning, long and short-term memory, audio-visual perception, and logical processing. These are thought to be set at a young age, and do not change throughout one's life. However, recent research has shown that cognitive skills are in fact developable (e.g., Kautz, Heckman, Diris, Ter Weel, & Borghans, 2014).

Non-cognitive skills, on the other hand, are skills that can be learned. Some of the most discussed non-cognitive skills in the literature include grit (Duckworth et al., 2007), motivation (Cofer & Appley, 1964; Atkinson, 1964; Deci, Vallerand, Pelletier, & Ryan, 1991), self-determination (Ryan & Deci, 2000; Deci, & Ryan, 2002, 2010; Gagné & Deci, 2005; Niemiec &

Ryan, 2009), optimism (Teahan, 1958, Scheier & Carver, 1985; Carver, Scheier, & Segerstrom, 2010), self-discipline (Duckworth & Seligman, 2005, 2006), hope (Snyder, 1994; Snyder et al., 1991), mindset (Dweck, 2006, 2015), and grit (Duckworth et al., 2007; Duckworth & Quinn, 2009). These skills are not fixed; they are able to fluctuate greatly during one's life.

Many psychologists have focused on cognitive skills in relation to student success and future financial outcomes. In a longitudinal study on cognitive skill distribution and economic growth, Hanushek & Woessmann (2012) showed that cognitive skill improvements in school systems lead to higher long-term growth of economies. Murnane, Willett, Duhaldeborde, and Tyler (2000) conducted a longitudinal study that aimed to find out whether or not cognitive skills can be used to predict future success in terms of financial earnings as well as their ability to complete college. They found that students with stronger basic high school cognitive skills earned, on average, more than those with weak cognitive skills. The results were significant because they found that a one-point difference in the math score of male high school seniors in 1982 is associated with a 1.5 percent difference in annual earnings at age 27. The financial earnings are also higher for those with more formal education. Economists Boissiere and Sabot (1985) found that reasoning ability had a low return in the labor market while literacy and numeracy had a high return. In other words, it would appear that literate and numerate workers are more productive. Farkas (2003) discussed the differences that cognitive and non-cognitive have on labor market success. He found that up to 80% of their labor market success could be contributed to non-cognitive skills, while only 20% could be contributed to cognitive skills. Carneiro, Crawford, and Goodman (2007) found that a person's IQ was not related to their level of persistence. They also found through regression models that social skills are very important

for schooling outcomes. Social skills were not directly related to all outcomes, such as literacy, but did have an influence on how much education they would complete in their lifetime.

There has been a profound movement toward investigating how non-cognitive skills can be applied to student success. As mentioned before, researchers have been developing scales to measure non-cognitive constructs in order to better assess student learning, capabilities, and success. Since non-cognitive skills are able to be developed, learning how to assess and develop them can aid educators to maximize students' efficiency in learning and ultimately their success in learning. For example, Tracey and Sedlacek (1987) found that the Non-cognitive Questionnaire (NCQ) was a valid predictor for college graduation rates in a White and Black student sample. The NCQ is created to measure contextual and experimental intelligences, and it is based on Sternberg's (1985) triarchic theory of intelligence. The NCQ consists mainly of self-reported item endorsements based on a five-point Likert scale, but it also includes opened-ended questions.

One non-cognitive skill discussed in the literature is self-discipline. Duckworth and Seligman (2005) found that factors related to self-discipline were stronger predictors for students underachieving than IQ scores. This would imply that the reasons students fail to succeed in their classes (and ultimately life) have more to do with non-cognitive factors than cognitive factors. This is actually great news for educators because students can learn these non-cognitive skills, whereas their intellectual ceiling is relatively fixed. Therefore, students can overcome their 'intellectual shortcomings' by being self-disciplined and motivated to learn. If teachers can infuse students with the love for learning and teach them how to be self-disciplined at a young age, they will have a higher likelihood of achieving success in their future.

Another non-cognitive skill that is popular and still relevant in the literature is self-efficacy. Self-efficacy has been examined by researchers for the past four decades and has been labelled as a highly effective predictor of students' motivation and learning. Bandura (1977, 1986, 1997) introduced self-efficacy as one of the key components in social cognitive theory and defines self-efficacy as personal judgements on one's capabilities to organize and execute courses of action to attain goals. In other words, self-efficacy is how well one believes in himself. Self-Efficacy Theory (SET) has become a major influence on research, education and clinical practice. In education, self-efficacy can be defined as the belief in one's capabilities to achieve a goal or an outcome. The more self-efficacy one has, the more likely they are to challenge themselves with difficult tasks. Likewise, students who have high self-efficacy are typically more intrinsically motivated (Bandura 1986). There are, though, cases where students with high levels of self-efficacy exhibit signs of lowered motivation. Jernigan (2004) found that over efficacious foreign language learners had lower motivation to study.

Mindset is another area of research related to a person's potential for future success. In social and behavioral research, entity theorists believe that ability levels are fixed, while incremental theorists believe that ability is malleable and therefore can be increased. Dweck (1999, 2006) described the ideas of fixed mindsets and growth mindsets as non-cognitive skills. Dweck argued that there are other factors besides intelligence and ability that act as determinants of students' classroom successes. A fixed mindset suggests that a person's intelligence is crystalized and is unable to change significantly over time. The inverse of this is the growth mindset, which suggests that intelligence is malleable and capable of significant change over time, typically through effort and motivation. For example, it is possible for young talented people to fail to achieve at levels consistent with their academic potential. Consequently, these

students may lower their academic expectations which could result in not attending college (Diener & Dweck, 1978, 1980), or perhaps dropping out prior to graduation (Hanson, 1994).

This phenomenon is described by Hanson as “lost talent.”

Not far from mindset, Snyder et al. (1991) have discussed the notion of adult hope and how hope plays a role in people’s motivation and future outcomes. Snyder (2005) defined hope as the process of thinking about one’s goals, along with the motivation to complete the goals. In addition to having optimism, a hopeful person makes positive things happen rather than simply waiting or wishing for things to happen. Snyder et al. (2002) extended the use of hope and academic success in college. In their six-year longitudinal study, the authors investigated if the Hope Scale scores correlated with student success. They found that the Hope Scale scores provided reliable predictors about college students’ academic performance over the course of their undergraduate careers in three areas; higher hope scores were related to higher cumulative GPA, higher likelihood of graduating, and lower likelihood of being dismissed due to poor grades.

Personality and behavioral traits have also been measured collectively rather than individually. The Big Five personality dimensions (John & Srivastava, 1999) is one such scale and has been used in research to predict a number of things, namely academic success and potential job performance. As the name suggests, there are five very broad personality dimensions: extraversion or surgency, agreeableness, conscientiousness, emotional stability versus neuroticism, and intellect or openness (Goldberg, 1990). Extraversion or surgency are related to how talkative, assertive, and energetic an individual is. Agreeableness is concerned with how good-natured, cooperative, and trustful an individual is. Conscientiousness is related to how responsible and dependable an individual is. Emotional stability versus neuroticism deals

with whether a person is calm, neurotic, and easily upset. Intellect or openness is related to a person's intellectual, imaginative, and independent-minded traits.

The Big Five personality dimensions have been used in the literature to determine success in academics. The conscientious dimension has been shown to be the most consistent of the dimensions in terms of being consistently associated with academic success (O'Connor & Paunonen, 2007). Conscientiousness has been shown to be a significant predictor of high school and college GPA, while openness is a significant predictor of SAT verbal scores (Nofle & Robins, 2007). Research has also revealed that conscientiousness can be a partial mediator of the relationship between motivation and college GPA (Komarraju, Karau, & Schmeck, 2009). The findings in the literature suggest that personality predictors (e.g. non-cognitive skills) account for the variance in academic performance beyond the variance accounted for by cognitive ability factors.

It is clear that previous researchers have found importance in non-cognitive skills as they relate to student success. This study therefore expands research in this area.

Conscientiousness and Goal Orientation

Conscientiousness and goal orientation have been exhaustively studied. Conscientiousness is a trait that involves being reliable, self-disciplined, hardworking, and persevering (McCrae & Costa, 1987). This construct has been linked to positive work outcomes and has been used as a predictor of job performance across a variety of different occupations (Barrick & Mount, 1991). Goal orientation is a dispositional variable that consists of two forms. The first form is developing new skills as a way to increase competence. The second form is a performance orientation where demonstrating competence is done by meeting a normative-based standard. Goal orientation stems from Dweck's early work with children (Dweck, 1986, 1989).

In this research, Dweck describes adaptive (and maladaptive) behavioral patterns that emerged on the basis of children's beliefs about ability (Dweck, Hong, & Chiu, 1993).

Colquitt and Simmering (1998) examined conscientiousness and goal orientation as predictors for motivation to learn, and as moderators of reactions to performance levels during the learning process. In this longitudinal study, 12 items from the NEO Five-Factor Inventory (Costa & McCrae, 1992) were used to measure conscientiousness and Button, Mathieu, and Zajac's (1996) eight-item dispositional goal orientation measures were used to assess goal orientation. The results showed that conscientiousness and learning orientation had a positive correlation with expectancy, valence, and motivation to learn. Performance orientation had a negative correlation with expectancy and motivation to learn. The results also showed that learners who are highly conscientious, highly learning oriented, and less performance oriented were significantly more likely to perform well on the Rosenthal and Rosnow's (1991) convention. These findings have suggested that people with higher levels of personality variables such as conscientiousness and goal orientation are more likely to have higher levels of motivation to learn. Thus, personality variables (or non-cognitive traits) are strongly related to one's motivation to learn.

Grit

The non-cognitive construct which is the focus of this study pertains to grit. Duckworth et al. (2007) defined grit as "trait-level perseverance and passion for long term goals" (p.166). The key difference between grit and determination is that grit implies that an obstacle needs to be overcome. Such obstacles could be a barrier challenge (something preventing you from achieving a goal), long-term goals, and challenging goals.

Duckworth and Quinn (2009) found that grit was a stronger predictor for achievement in challenging domains than measures of talent. They chose four different sample groups, and each one yielded the same results. The grit test that the researchers created was a stronger predictor for success than the Big Five dimensions, education, age, gender, and career changes. The data yielded significant results equally for both male and female participants.

However, this construct has seen a fair amount of controversy in the literature. Researchers have debated the validity and predictability of grit as Duckworth defined it. Strayhorn (2014) discussed the implications of grit for African American students in the context of a predominantly white higher-education arena. Arthur, Kristjánsson, and Thoma (2016) highlighted issues of grit as a valid construct in relation to how Duckworth defined grit in her book *GRIT: The power of passion and perseverance* (Duckworth, 2016). A key issue is how the 10-item self-report survey is, to use Duckworth's own words, "ridiculously fakeable" (p. 229). In the initial description of grit as a psychological construct, grit was composed of both resilience and persistence (Duckworth et al., 2007). However, these concepts are downplayed in her book, and rather than focusing on resilience, she emphasized passion.

So why grit? Grit as a construct has been widely studied in the field of psychology. It is a construct that has potential to be further researched in the future. However, what researchers have not agreed upon is how to properly assess grit, and if grit plays a role in success across different populations (e.g., people from different demographics or SES). While Duckworth's Grit-S scale provides a quick method to assess grit, it has the same shortcomings as the previously mentioned scales; it is prone to manipulation and response bias, and it relies on simple item endorsement without context. I believe that the development and validation of a

vignette-based grit scale can significantly contribute to future studies on measurement of non-cognitive skills.

Since research has shown that grit and student success are highly related, this study will investigate this relationship as well as create a new vignette-based proxy measure for an attribute of academic success. This particular set of variables has been examined in the past, thus this study should confirm and strengthen the argument that a person's level of grit is connected to their overall scholastic performance.

Grit, Race, and SES

What is contentious in the literature relates to grit and success regarding a person's ethnicity. Strayhorn (2014) found that Black male collegians with higher levels of grit achieved higher grades in post-secondary education. Gorman (2015) found no statistical differences in race in regard to student success measured by self-reported grades, but age was a significant factor. However, qualitative inquires at KIPP charter schools suggest that grit in and of itself is not enough to become successful; it is entirely possible to be "gritty" and still be unsuccessful. The purpose of this study was not to concretely state that grit will always lead to success. Rather, that grit plays a role in becoming successful, specifically in the context of education. Additionally, no previous research has been done to confirm Duckworth's theory that individuals from low SES exhibit more grit than those from higher SES backgrounds (Duckworth, 2016). This research aimed to address this gap in the literature and investigate if SES is a moderator variable of grit.

Exploring the Role of SES in Educational Assessments

Since the relationship between grit and academic achievement with SES as a moderator variable is being investigated, the study also had to consider how to best measure SES. There has

been much debate on how to properly and accurately assess SES in education assessments. In the literature, SES has been measured by different variables (Sirin, 2005) and thus has been difficult to define what exactly SES includes. Baer, Baldi, and Merola (2005) recommended to the National Center for Education Statistics (NCES) to consider three theoretical concepts for measuring SES: educational attainment, occupational status, and financial resources. For some historic context, the first conceptualization of SES and academic achievement was by Taussig (1920), and this conception was based entirely on the father's occupational status. Over time researchers have added things when assessing a students' SES, such as home possessions (Sims, 1927) and cultural and material possessions (Chapin, 1933). Today, SES is being assessed in students on a worldwide scale. For example, the Organisation for Economic Cooperation and Development (OECD) conducted the Programme for International Student Assessment (PISA), and students were given items based on three variables: an index of home possessions, educational attainment of the parent, and occupational status of the parent. For the purposes of this study, recommendations by NCES were used, including the primary components of SES "big three" variables: family income, parental educational attainment, and parental occupational status.

Concerns with Traditional Self-report Assessments

The aforementioned scales have both strengths and weaknesses. First, the vast majority of these assessment tools are self-reported. The inherent advantage of using self-reported assessments is that they are typically inexpensive to administer. They also provided the researcher with an efficient way to measure the construct. This is more so the case with scales that measure behavior and non-cognitive skills, as there are typically fewer items than those found in cognitive assessment.

However, there are two glaring issues with self-report assessment tools. The first is they are susceptible to response bias (Fan et al., 2006). As they are self-reported, opposed to observed, it is entirely possible to provide false responses, whether consciously or subconsciously. This would lead to an invalid measure of the construct. Another pitfall is that data are simply comprised of item-endorsements. In other words, the response only requires the examinee to endorse an item, typically on a Likert scale, with little to no context. For example, let us consider a Likert scale with response choices ranging from *strongly disagree* to *strongly agree*. Without context or a detailed explanation, it is likely that *strongly agree* for one person may only be *agree* to another. In other words, without context, it is entirely possible for examinees to interpret and respond to the items differently, which would lead to an invalid measurement of the construct.

Self-administered questionnaires (SAQ) have proved popular in research because this format can provide anonymity and privacy to the respondents and thus allows respondents to answer more candidly on sensitive issues (Babbie, 1995). This anonymity and privacy can also reduce the likelihood to produce socially desirable responses when compared to face-to-face interviews (Fujii, Hennessy, & Mark, 1985). SAQs are typically much less expensive than other methods of data collection (e.g. face-to-face interviews) because they do not require as many external resources (e.g. trained interviewer).

However, the many pitfalls of SAQs cannot be ignored. Alexander and Becker (1978) discuss the shortcomings of questionnaires and interviews in studying human attitudes and behaviors because they fall victim to bias self-reports. In the same vein, Austin, Deary, Gibson, Murray, McGregor, and Dent (1998) investigated the individual response spread in self-report scales. Their results showed that in their sample, there was a statistically significant spread in

responses (low test-retest correlations) which can be viewed as a validity threat. Additionally, Duckworth and Yeager (2015) highlight the need for other forms of measurement to assess behavioral and personality traits, including the Grit-S scale. Researchers have also suggested that supplementing self-report questionnaires with anchoring vignettes can reduce reference bias (King, Murray, Salomon, & Tandon, 2004; Kyllonen & Bertling, 2013).

Fan et al. (2006) also explored the inaccuracy and invalidity in self-report surveys. By using Add Health data, the researchers were able to provide evidence that some adolescents gave either inaccurate or invalid responses to a self-administered questionnaire. The researchers distinguished between inaccurate responses (inaccurate or false responses due to carelessness or confusion) and jokesters (intentional false responses). Their findings showed that on psychological and behavioral outcome variables, the distortion effects were significantly greater in jokesters than inaccurate respondents. Their conclusion was this effect could pose a serious challenge for validity when focusing on special subgroups.

Vignettes over Traditional Item Endorsement in Measurement

There are ways to make improvements on the traditional item-endorsement method of scale development. Some researchers advocate the use of mapping sentences to create more valid items (Shye, Elizur, & Hoffman, 1994; Borg & Shye, 1995; Beghtol, 1995; Randall & Engelhard, 2010). The sense of mapping sentences is creating items with varying levels of the established facets of a construct. For example, consider a researcher who is developing a scale to measure motivation, and the literature suggests that motivation has two facets. The researcher decides that each facet has two levels (e.g. high and low). The researcher would write items that vary in the different levels for each facet until all combinations have been covered.

Another approach, and the approach used in this research, is the development of vignettes. A vignette is a short story that presents an issue and has no ending. The use of vignettes in measurement has been well-documented in the literature. In the realm of psychological and sociological research, a vignette is an anecdotal hypothetical situation that participants respond to as a method to reveal their perceptions or impressions of events. These can be in the form of a longer narrative or perhaps in shorter form (e.g., a few sentences). Nosanchuk (1972) began the discussion of using vignettes as an experimental approach to the study of social status. The researcher proposed the use of vignettes as an experimental practice. Alexander and Becker (1978) described the benefits of vignettes in survey research, stating that vignettes can produce more valid and reliable measures than shorter, simpler items that are used in standard survey research practice due to the added context.

The medical and health fields have been using vignettes in assessment for years (Dowd & Todd, 2011; Peabody, Luck, Glassman, Jain, Hansen, Spell, & Lee, 2004; Peabody, Luck, Glassman, Dresselhaus, & Lee, 2000; Norcini, 2004; Dresselhaus, Peabody, Luck, & Bertenthal, 2004; Bago d'Uva, Van Doorslaer, Lindeboom, & O'donnell, 2008; Hughes & Huby, 2002). Additionally, the use of vignettes has been used to assess self-rated health (Salomon, Tandon, & Murray, 2004; Grol-Prokopczyk, Freese, & Hauser, 2011; Brondani, MacEntee, Bryant, & O'Neill, 2008).

However, in the fields of educational psychology, measurement, and education, specifically assessing non-cognitive skills, the use of vignettes can be found but is not as widespread. One method of item-writing based on vignettes is intermediate concept measure, or ICM (Bebeau & Thoma, 1999) which has been used for measuring moral decision making. In education, vignettes have been used mainly as a learning tool rather than a form of evaluation or

assessment (Roach & Wandersee, 1993; Berque, Prey, & Reed, 2006; Brown, 2000; Chau, Chang, Lee, Ip, Lee, & Wootton, 2001; Spalding, 2004). Vignette-based surveys have also been used to show application of Rasch-based statistical analyses (Martin, Campanelli, & Fay, 1991; Gumpel & Wilson, 1996; Campbell, 1996). While these studies utilize a vignette-based approach, this study is one of the first to use vignettes to measure a non-cognitive construct.

The development of vignettes, and scale development in general, can greatly benefit from focus groups. Focus groups are small diverse groups of people who participate in a guided discussion (Strayhorn, 2014). Focus groups can be structured, semiconstructed, or unstructured in nature. Focus groups are popular in a number of fields, namely qualitative marketing (e.g., Calder, 1977), but have also been used in medicine (e.g., MacIntosh, 1993; Basch, 1987; Wilkinson, S. (1998) and psychology (e.g., Stoll-Kleemann, O’Riordan, & Jaeger, 2001). In scale development, focus groups can assist the researcher in a number of ways, including defining facets of a construct, writing items based on anecdotal evidence collected from the participants, and even helping the researcher check face validity of an assessment tool. Since focus groups are meant to have a diverse sample of participants, it provides the researcher with qualitative data related to the target population. In other words, focus groups can provide insight on the construct the researcher is interested in creating a measurement tool for, and shed light on issues that the researcher may have overlooked.

While the use of vignettes is not a perfect form of assessment, this study should shed light on whether the vignette-based approach provides a more valid form of measurement. Admittedly, observable data may prove more useful when considering the possibility that examinees provide falsified responses. Where vignettes excel is the ability to provide context to both the items and responses. Opposed to endorsing an item out of context and providing a

response on a Likert scale, examinees would be provided a scenario and asked how they would respond (or how the person in the vignette should respond). To get the most out of focus groups, the researcher can conduct an initial focus group to gather information, then conduct follow-ups to check if their interpretation of the data collected from the focus groups closely matches those of the participants.

The aim of this study was to address the issues researchers have discussed concerning grit measurement, as well as use vignettes to develop a more powerful, and thus predictive, scale than the current grit scale (Duckworth & Quinn, 2009). Grit is going to remain the same construct as defined by Duckworth et al. (2007) but the items will be written in the context of education. The research questions for this research were:

1. What are the psychometric properties of the Grit-V scale;
2. Does the Grit-V scale better relate to success in academia than the Grit-S scale and Big Five dimensions;
3. Does the use of vignettes provide stronger evidence of validity when measuring grit compared to short sentence items; and
4. Is SES a moderator variable when determining a person's success in relation to their level of grit?

CHAPTER III:

METHODS

The procedure used to develop this new vignette-based assessment required a multi-step process. First, semi-structured focus groups were conducted as a way to inform the perspective on grit. This resulted in themes and participant anecdotes related to grit for the study. Second, after exploring the focus group data, the item pool for the pilot study was created that employed all of the components of grit that came out of the focus groups. The pilot study was the first time these items were used, and the purpose was a preliminary test of the item pool. During the pilot study, items were investigated that provided the most information. Once the final pool of items was decided upon, the final scale, along with the Grit-S scale, Big-Five personality scale, AMS-C, NSSE, and demographic information survey was prepared and distributed to the sample population via Qualtrics. The statistical analyses run on the final data included Many-facet Rasch analysis, higher-order bifactor model via confirmatory factor analysis (CFA), an exploratory factor analysis (EFA), multiple regression, and binary logistic regression. The many-faceted Rasch analysis provided item and participant calibrations, as well as fit indices. CFA was used to check the factor structure of the Grit-V. EFA was used to explore other possible factor structures that may produce a better model fit. The multiple regression analysis provided information on the relationship between grit and academic success, as well as helped determine which grit scale provided a better predictor for academic success. Logistic regression was used to investigate the relationships between grit and classes dropped and changed majors variables.

Step 1: Focus Groups

Participants

The participants ($N=15$) took part in one of the three focus groups ($n=5$) and consisted of the target population for which the scale would be used (university undergraduate students). The sample size of the focus groups was chosen to balance the amount of information gathered and logistics of conducting the focus group; a sample size of five provided sufficient data while reducing the chances of people speaking over each other, in turn easing the data collection and future categorization process (Kamberelis & Dimitriadis, 2005; Barbour, 2007). The participants were from undergraduate programs in different subject disciplines (e.g., math, geology, psychology, education, and kinesiology). Additionally, to further consider the target population for the scale, there was a balance of participants from different gender groups and different racial backgrounds in each of the three focus groups. Table 3.1 contains the demographic information from the focus groups.

Table 3.1

Demographic Information for Focus Groups

	Number
Gender	
Male	6
Female	9
Age	
Mean	20.1
Median	19
Mode	20
Range	18 - 31
Education Level of Parents	
Did not complete HS	0
High School	1
Some College	2
Associate Degree	1
Undergraduate Degree	10
Graduate Degree	1
Race/Ethnicity	
White	5
Hispanic	4
Black or African American	3
Asian or Pacific Islander	2
Native or American Indian	0
Other	1

Data Collection

Data were collected by conducting three semi-structured focus groups (see Appendix A for semi-structure outline). The participants ($N=15$) were recruited via flyers posted around The University of Alabama campus and were selected based on their availability. A schedule was sent to participants who were interested in participating, and the ones available during those times were selected. Each focus group meeting was approximately one hour in length and held in an available classroom at The University of Alabama. The focus groups were audio recorded with a

digital audio-only voice recorder. After each focus group was completed, the data was transcribed verbatim in order to be coded. Follow-ups with participants were conducted by either phone or email to ensure that the information gathered from the focus groups was accurate and complete.

Instrument

As this stage in the study is a qualitative inquiry, it is important to consider myself as an instrument of the research (Lincoln & Guba, 1985). In this case, it is appropriate to provide details on my background and perspective so that the reader can better interpret my understanding of the data (Elliot, Fischer, & Rennie, 1999). I have been doing research on grit for five years, and therefore have been exposed to the most current and relevant research concerning grit. To avoid biasing the focus group and the findings, I strictly adhered to the focus group guidelines (see Appendix A). I collected all the data and was guided under the supervision of a faculty with extensive formal training in focus group moderation. I have experience working with the targeted population, collecting qualitative data, and conducting qualitative research.

Procedure

The purpose of the focus groups was to gather qualitative information on the construct grit from the participants. Information included anecdotes from participants that exemplify grit in an authentic academic context, as well as their experiences with long-term goals and how they persevered through obstacles. This was meant to provide authentic descriptions and experiences of grit, and the data gathered assisted me in writing the initial item pool.

The three focus groups took place over a one-month period and were held in a classroom at The University of Alabama. Before each focus group session, a series of ground rules were developed and agreed upon by the participants and I (e.g., not talking over each other or

interrupting each other or making inappropriate comments) to facilitate mutual trust and respect. The participants were informed of the audio-recording process. Participants were asked to participate in a follow-up focus group to discuss the initial item pool. The set of items for the semi-structured interview were asked in order (see Appendix A) with development of follow-up questions dependent upon the responses of the participants.

Once the focus group audio recordings were transcribed, I used the computer software package QSR Nvivo 11 to assist with the data management, as well as sorting and retrieval. I undertook an inductive content analysis of the data in a three-step process. First, the data were reviewed as a whole to become familiar with the data and then undertaken to identify meaningful units (e.g., words, phrases, sentences, and paragraphs). Next, the meaning units were clustered in order to build themes in the data. The purpose was to identify the participants' perspectives of the definition of grit, as well as the different facets that describe grit as a construct. Last, experts in the field of educational and behavioral psychology ($N=5$) reviewed the themes, and the themes were modified accordingly. The final themes were used in the item development process of the Grit-V scale.

Step 2: Item and Scale Development

The data used to facilitate the development of items were the themes derived from the focus group transcriptions. Upon reviewing the themes composed from the focus group data, I began writing items following the guidelines and suggestions set by previous scholars in item development (DeVellis, 2016; Downing & Haladyna, 1997; Haladyna & Downing, 2011; Haladyna & Rodriguez, 2013) and vignettes (Alexander & Becker, 1978; Nosanchuk, 1972; Randall & Engelhard, 2010).

First, a pool of items ($N=15$) were written that captured the notions of the facets of grit, which were perseverance of effort and consistency of interest (see Appendix B). For consistency, half of the items captured the facet perseverance of effort, and half captured the facet consistency of interest. In addition to items measuring grit, there was a check question that served the purpose of ensuring participants were providing valid responses to the items.

In order to check the content validity, a panel of experts ($N=5$) whose background is in educational psychology, behavioral psychology, research methodology, and measurement reviewed the items and provided feedback. Following in the footsteps of previous researchers (e.g., Davis, 1992; Tilden, Nelson, & May, 1990), the panel of experts were given the initial item pool and provided a survey for each item to assess the items' relevancy to the construct on a four-point Likert scale. The responses were *1 – relevant to the construct*, *2 – relevant but needs minor alteration*, *3 – unable to assess relevance without revision of item*, and *4 – not relevant to the construct*. The reviewers checked the items independently and met with me to discuss problem items or items that they deemed insufficiently relevant for the scale. This process continued until there was agreement on the majority of items.

Step 3: Pilot Study and Preliminary Analysis

Procedure

Once a pool of items was written, I conducted a pilot study to identify both strong and weak items in terms of their factor loading through confirmatory factor analysis, as well as see if there was a positive correlation between scores on the Grit-V and the Grit-S scale. Once the results from the pilot study were collected, I analyzed the results and decided which items to revise and keep for the final version of the scale. The revision process continued until the panel

of experts, focus group members, and I agreed that the final item pool was representative and captured the facets pertaining to grit as a non-cognitive trait.

Participants

The participants represented the target population (university undergraduate students, $N=20$). The participants ranged in age from 18 – 39 years with a mean age of 20.5 and a mode of 21. The pilot included students from undergraduate programs in different subject disciplines (e.g., math, geology, psychology, education, and kinesiology). See Table 3.2.

Table 3.2

Demographic Information for Pilot Study

	Number
Gender	
Male	7
Female	13
Age	20.5
Mean	20
Median	21
Mode	18 - 39
Range	
Education Level of Parents	0
Did not complete HS	1
High School	4
Some College	2
Associate Degree	11
Undergraduate Degree	3
Graduate Degree	
Race/Ethnicity	
White	9
Hispanic	2
Black or African American	4
Asian or Pacific Islander	3
Native or American Indian	0
Other	2

Data Analysis

Partial credit Rasch model analysis. The purpose of this analysis was to examine the calibrations of the items and participants to determine which items were functioning properly as well as check if there was an acceptable spread of item difficulty. Fit statistics were examined to determine if the data fit the Many Facet Rasch model. Additionally, a differential item functioning (DIF) analysis was conducted to investigate whether the items function equally across gender, racial, and different SES backgrounds by investigating the interactions between these facets.

A preliminary Partial Credit Rasch model (Andrich, 1978; Linacre, 1989) was run to explore the calibrations of items and participants, as well as investigate the fit statistics for the initial item pool. Rasch software Facets version 3.71.4 (Linacre, 2015) was used to conduct the model analysis. The model consisted of five facets: participants, items, gender, race/ethnicity, and SES. The equation for the model was as follows:

$$\ln \left[\frac{P(nij)(m=x)}{P(nij)(m=x-1)} \right] = \Theta_n - \delta_i - \gamma_j - \pi_k - \lambda_l - \tau_m$$

where Θ_n is the logit-scale location for participant n ; δ_i is the logit-scale location for item i ; γ_j is the logit-scale location for gender j ; π_k is the logit-scale location for race k ; λ_l is the logit-scale location for SES l ; and τ_m is the location on the logit-scale where rating scale categories k and $k-1$ are equally probably across participants, items, gender, race, and SES.

Confirmatory and exploratory factor analysis. A confirmatory factor analysis was run using SAS version 9.4. My goal was to identify the most parsimonious factor structure by investigating the fit indices (e.g., RMSEA, GFI, and CFI values). Whether the items were loading correctly to the appropriate factors was also checked. Next, an exploratory factor analysis was run in SPSS version 25 to evaluate the factor structure of the Grit-V scale to explore

if a more parsimonious model existed. The most parsimonious factor solution was chosen based on the eigenvalues (factors with eigenvalues less than 1.0 were omitted). This process facilitated the decision-making process on which items needing to be revised or removed from the final form of the scale.

Validity. To check construct validity (in this case, factorial validity), a confirmatory factor analysis (CFA) model to confirm the factor structure during the pilot study was conducted. If there were any issues with the factor loadings, those items were revisited and rewritten; or items were eliminated that did not fit either statistically or theoretically in the measurement model.

Convergent validity (the degree to which two measures of constructs that theoretically should be positively related) was also assessed (Wilson, 2005). Von Culin, Tsukayama, and Duckworth (2014) noted that motivation and engagement were positive correlates of grit; therefore, I expected to find similar relationships when conducting the correlation study between grit, motivation, and engagement. The Academic Motivation Scale – College Version (AMS-C) (Vallerand, Pelletier, Blais, Briere, Senecal, & Vallieres, 1992) and the National Survey of Student Engagement (NSSE) were used to identify convergent validity. In Duckworth and Quinn's (2009) original work, the researchers compared the grit scale with the Big Five personality dimensions (John & Srivastava, 1999) in terms of their ability to predict future success. Since they found that the grit scale was more predictive, it was expected in this study that similar results would be found with the Grit-V scale.

Therefore, a correlation study was conducted comparing scores from the Grit-V and Grit-S, NSSE, BFI dimensions, and AMS-C via SPSS version 25. If the Pearson correlation coefficients are both significant and positive for the Grit-V, NSSE, AMS-C, and BFI dimensions, that would provide evidence of convergent validity. If the Grit-V and Grit-S have a statistically

significant positive Pearson correlation coefficient and similar monotonic trend, this would signify that the two scales are measuring grit similarly (Field, 2013).

Revisions and final scale form. Once the pilot study was conducted, and the data collected and analyzed, a decision was made about which items needed to be revised or removed from the final form of the scale. Once the revisions were made, the edited items were presented to the panel of experts independently who reevaluated the items. This process continued until the panel of experts and I were in agreement on the final form of the scale. Additionally, six follow-up phone calls with focus group members were made to share the final item pool with the participants.

Step 4: Final Item Pool and Analyses

Participants

The participants ($N=240$) were undergraduates at The University of Alabama. A post hoc power analysis via G*Power 3.1 revealed that a sample size of $N=240$ provided a power ($1 - \beta$ error probability) of 0.999 and effect size $f^2 = 0.15$ with three tested predictors and seven total predictors. The sample of participants resembled the target population of the university, which includes a diversity of racial, SES, and gender backgrounds. See Table 3.3 for demographic information for the final analysis. Only undergraduate students in their junior or senior year were used in the analysis. This was to ensure students had taken enough coursework to properly assess their academic success.

Table 3.3

Demographic Information for Final Analysis

	Number
Gender	
Male	97
Female	143
Age	
Mean	20.1
Median	20
Mode	20
Range	18 - 34
Education Level of Parents	
Did not complete HS	9
High School	10
Some College	37
Associate Degree	29
Undergraduate Degree	101
Graduate Degree	54
Race/Ethnicity	
White	107
Hispanic	27
Black or African American	51
Asian or Pacific Islander	34
Native or American Indian	9
Other	12

Measures and Instruments

Grit. Two different grit measures were used: Grit-V and Grit-S scales. The measure grit was assessed using the Grit-V scale (see Appendix C). The final form had eight items designed to measure grit in the context of academia, and one validity check response. The responses were in Likert scale form and had four response levels.

The measure grit was also assessed using the Grit-S scale (see Appendix D). The Grit-S scale contained ten items, and all responses were assessed using a five-point Likert scale ranging from *not like me at all* (1) to *very much like me* (5).

Motivation. The measure motivation was assessed using the Academic Motivation Scale (AMS-C). The AMS-C consisted of 28 items, and the responses were assessed on a seven-point Likert scale form, ranging from *does not correspond at all* (1) to *corresponds exactly* (7). The internal consistency of response for the subscales, assessed with Cronbach alpha, varied from .83 to .86. This was checked to verify that the instrument was working as intended. See Appendix E for more information on the AMS-C.

Student engagement. The measure student engagement was assessed using a shortened version of the National Survey of Student Engagement (NSSE). The NSSE is a 40-item survey used to assess college-level student engagement. Item responses vary from Likert scale to open-ended. Items 1 – 3, 6, 8, 9, 15, and 16 ($n=8$) were used with permission from *The College Student Report*, National Survey of Student Engagement, Copyright 2001-18 The Trustees of Indiana University. These items are specifically related to the student, while the others are related to the instructors and course content. See Appendix F for more information on the NSSE.

Academic success. The measure academic success was measured on the academic success and demographic information questionnaire (further referred to as the academic success questionnaire, ASQ) which I created. Academic success was measured by cumulative GPA (continuous), whether students dropped or withdrew from classes due to difficulty (dichotomous), and whether students changed majors (dichotomous). See Appendix G for more information on the ASQ.

Personality dimensions. The measure personality dimensions were measured using the Big Five Inventory (BFI) (John & Srivastava, 1999). The BFI is a 44-item questionnaire that measures an individual on the Big Five Factors of personality (Goldberg, 1990, 1993). The five factors are *extraversion*, *agreeableness*, *conscientiousness*, *neuroticism*, and *openness*. Item responses are assessed on a five-point Likert scale ranging from *disagree strongly* (1) to *agree strongly* (5). The average Cronbach alpha value was .83. I checked this to verify that the instrument was working as intended. See Appendix H for more information on the BFI.

Socioeconomic status. The measure SES was measured by assessing the primary components of SES “big three” variables: family income, parental educational attainment, and parental occupational status. A question for each variable was included in the parent information section of the ASQ (questions 7 – 9). The data was categorical. Parental occupational status was a fill in the blank question, and the data was recoded to fit into a categorical structure on a five-point ordinal scale.

Procedure

The data were collected using convenience sampling. Flyers were distributed around The University of Alabama campus calling for volunteers to participate in the study. Once the participants contacted me, an email invitation was sent that contained a link to the survey. Participants were given the final form of the Grit-V scale, the Grit-S scale (Duckworth & Quinn, 2009), the ASQ, BFI (John & Srivastava, 1999), AMS-C (Vallerand, Pelletier, Blais, Briere, Senecal, & Vallieres, 1992) and the NSSE. The aforementioned surveys were available via the online survey platform Qualtrics, and the College of Education at The University of Alabama assisted in the uploading of the scales to the Qualtrics platform.

Analysis

Research Question 1

What are the psychometric properties of the Grit-V scale? For research question one, the final form of the Grit-V scale was analyzed for its psychometric properties. The final form has ten items designed to measure grit in the context of academia. The final data was gathered from the participant ($N=240$) responses from the final form of the Grit-V scale. A sample of 240 was sufficient for Rasch and confirmatory factor analysis (Schumacker, 2015).

The psychometric properties were collected by conducting six statistical methods. The first two were used to determine the reliability of the scale (Cronbach's alpha and split-half method). The third was a Many Facet Rasch Model. This was used to explore the item and participant calibrations, as well as the fit indices. The fourth was confirmatory factor analysis. CFA was used to confirm the factor structure of Grit-V scale. The fifth was an exploratory factor analysis. EFA was used to check if there was a more parsimonious model. The final was an analysis of variance (ANOVA). This was conducted to check whether there were meaningful differences in the different levels of grit.

To check the reliability of the Grit-V, two analyses were conducted: Cronbach's alpha and split-half method. In these analyses, the Cronbach's alpha and Spearman-Brown coefficient were computed using SPSS version 25. A Cronbach's alpha and Spearman-Brown coefficient of .7 or above would suggest the scale had an adequate level of reliability (Miller and Lovler, 2016).

The model consisted of the same five facets used in the pilot study: participants, items, gender, race, and SES. The equation for the model was the same as the one used in the pilot study.

Similar to the preliminary analysis, the calibrations of the items were checked to determine which items were functioning properly, as well as, to determine if there was an acceptable spread of item difficulty. Fit statistics were examined to determine if the data fit the Many Facet Rasch model. Additionally, a DIF analysis was conducted to investigate whether the items functioned equally across gender, racial, and different SES backgrounds by investigating the interactions between these facets.

Similar to the preliminary analysis in the pilot study, a confirmatory factor analysis was run using SAS version 9.4. The goal was to identify the most parsimonious factor structure by investigating the fit indices (e.g., RMSEA and GFI values) of the final form of the Grit-V scale. Items were checked to determine if they loaded correctly on the appropriate factors. Next, an exploratory factor analysis was run in SPSS version 25 to evaluate the factor structure of the Grit-V scale. The most parsimonious factor solution was chosen based on the eigenvalues (factors with eigenvalues less than 1.0 were omitted).

To determine if there were statistically significant differences in the four levels of grit, an ANOVA model was run via SPSS version 25. The dependent variable was the Grit-V total score, and the independent variable was the grouping of the total Grit-V scores recoded into the four level scores (low, medium, high, very high). The responses were on an ordinal scale, with each response relating to specific level of the construct; the students selected a response that endorsed a specific level of grit. The range was based on the endorsements of these levels: students who endorsed primarily low grit responses (scores from 10 to 14) had low levels of grit; students who endorsed primarily medium grit responses (scores 15 to 24) had levels medium levels of grit; students who endorsed primarily high grit responses (scores 25 to 34) had high levels of grit; and students who endorsed very high grit responses (score above 34) had very high levels of grit.

Finally, since the Grit-V contained longer items and likely took longer to read and answer than the Grit-S, it was important to compare the actual time-on-task differences between these two assessments. To compare the average length of time students used to complete the Grit-V and Grit-S, the time stamps (start and end times) provided by Qualtrics were used.

Research Question 2

Does the Grit-V scale better relate to success in academia than the Grit-S scale, AMS-C, BFI, and NSSE? For research question two, the Grit-V, Grit-S, AMS-C, BFI, the reduced form of the NSSE scale, and ASQ were used. The data were gathered from participant ($N=240$) responses from the Grit-V, Grit-S, AMS-C, BFI, a reduced form of the NSSE scale, and ASQ.

The purpose of this analysis was to compare the Grit-V and the Grit-S scales and determine which had more predictive power in terms of academic success. Twelve analyses with six comparisons were run using SPSS version 25. See Table 3.4 for the summary of the analysis models. The assumptions of multicollinearity, homoscedasticity, and linear relationship of the variables were checked prior to the analyses (Fields, 2013). Additionally, Cronbach alphas were computed for the AMS-C and BFI to confirm that the instruments are working as intended (reliable internal consistency of responses).

To make meaningful comparisons between the predictive strength of grit, comparisons between the different regression models were conducted using statistics software SPSS version 25. The first comparison was between models A and B (simple regression). The rationale was to compare the Grit-V and Grit-S scales with no other variables to reduce the variance from the other factors. The second comparison was between models C and D (binary logistic regression) where classes dropped was the dependent variable and Grit-V and Grit-S score were the independent variables. The third comparison was between models E and F (binary logistic

regression) where changed majors is the dependent variable and Grit-V and Grit-S score are the independent variables. The fourth comparison was between models G and H (multiple regression). These models include Grit-V, Grit-S, personality dimensions, motivation, and engagement as independent variables, and GPA as the dependent variable. This revealed which variables have the strongest predictive strength in the regression equation. The fifth comparison was between models I and J (binary logistic regression), where classes dropped was the dependent variable and Grit-V, Grit-S, personality dimensions, motivation, and engagement as independent variables. The final comparison was between models K and L (binary logistic regression), where changed majors was the dependent variable and Grit-V, Grit-S, personality dimensions, motivation, and engagement as independent variables.

Table 3.4

Summary of Regression Models

Comparison	Model	Dependent Variable	Independent Variable(s)
1	A	GPA	Grit-V scale
	B	GPA	Grit-S scale
2	C	Classes dropped	Grit-V scale
	D	Classes dropped	Grit-S scale
3	E	Changed majors	Grit-V scale
	F	Changed majors	Grit-S scale
4	G	GPA	Grit-V scale Personality dimensions Motivation Engagement
	H	GPA	Grit-S scale Personality dimensions Motivation Engagement
5	I	Classes dropped	Grit-V scale Personality dimensions Motivation Engagement
	J	Classes dropped	Grit-S scale Personality dimensions Motivation Engagement
6	K	Changed majors	Grit-V scale Personality dimensions Motivation Engagement
	L	Changed majors	Grit-S scale Personality dimensions Motivation Engagement

For the final analysis, a MANOVA was used to explore if there were mean differences between the four levels of grit (low, medium, high, very high). The independent variables were GPA, student engagement, academic motivation, and the Big Five personality dimensions. Tukey post-hoc tests were run to investigate the mean differences between the grit levels and the independent variables. SPSS version 25 was used to conduct this analysis.

Research Question 3

Does the use of vignettes provide stronger evidence of validity when measuring grit compared to short sentence items? For research question three, the Grit-V, Grit-S, AMS-C, a reduced form of the NSSE scale, and ASQ was used. The data was gathered from participant ($N=240$) responses from the Grit-V, Grit-S, AMS-C, a reduced form of the NSSE scale, and ASQ. The data also included verbatim transcripts from the follow-up interviews.

To develop a strong and valid form of measurement, four forms of validity evidence were identified: criterion validity, construct validity, convergent validity, and content validity.

Criterion validity is the extent to which a construct is related to an outcome (AERA, APA, & NCME, 2014). The type of criterion validity was predictive validity; predictive validity is the extent to which a score on a scale predicts future outcomes. Evidence of predictive validity can be obtained by analyzing the relationship between scores on the assessment and a future outcome. In this case, the outcome (or latent variable) was academic success, which was measured using students' cumulative GPA, if they dropped or withdrew from classes due to difficulty, and if the student changed majors. It was expected that there would be a positive correlation between level of grit and GPA, and negative correlations between grit and classes dropped and changed major. A regression model was conducted to explore the relationships between grit (Grit-V scale), the academic success variables, motivation (AMS-C), and student

engagement (NSSE). The statistical software SPSS version 25 was used to analyze the relationship between the variables grit, motivation, and student engagement. The analysis from research question two was used to determine if there was sufficient evidence for criterion validity.

Construct validity is the extent to which the assessment tool is measuring the intended construct. Evidence of construct validity can be obtained by analyzing the relationship between the content of the assessment tool and the construct it is intended to measure (AERA, APA, & NCME, 2014). To check for evidence of content validity, the internal structure of the Grit-V scale was investigated. The dimensionality of the Grit-V was assessed via principal component analysis. Lastly, the Cronbach alpha level was checked for internal reliability response consistency. The statistical software packages Facets 3.80.4, SPSS version 25, and SAS version 9.4 were used for these analyses.

Convergent validity is the extent to which two (or more) measures of constructs are theoretically related. Evidence of construct validity can be obtained by analyzing the relationship between two measures (AERA, APA, & NCME, 2014). To gather evidence of convergent validity, a correlation analysis was conducted to investigate the relationship between scores on the Grit-V, NSSE, BFI dimensions, and AMS-C. SPSS version 25 was used to conduct this analysis.

Content validity is the extent to which the assessment actually measures the intended construct. Evidence of content validity was obtained through expert panel review (AERA, APA, & NCME, 2014). To provide evidence of content validity, an expert panel review was used in the item development process. The intraclass correlation for the initial item pool of the Grit-V

was computed during the pilot study process and was used to determine if there is evidence of content validity. SPSS version 25 was used to conduct this analysis.

Research Question 4

Is SES a moderator variable when determining a person's success in relation to their level of grit? For research question three, the Grit-V and ASQ were used. The data was gathered from participant ($N=240$) responses from the Grit-V and ASQ. The purpose of this analysis was to determine if SES is a statistically significant moderator variable in the model. A moderator regression analysis was run in SPSS version 25 to investigate the relationship between the dependent variable (GPA), and independent variable (grit) with SES as a moderator variable. Three models were run; each model used a different SES variable (parents' income, parents' education level, and parents' occupational status). A model summary is shown in Table 3.5. ANOVA models were run to test the mean difference in grit score by SES, where the total score from the Grit-V was the dependent variable and the SES variables were the independent variables.

Table 3.5

Model Summary for Moderator Analysis

Model	Dependent Variable	Independent Variable	Moderator Variable
M	GPA	Grit-V	SES (Parents' income)
N	GPA	Grit-V	SES (Parents' education level)
O	GPA	Grit-V	(SES)Parents' occupational status

CHAPTER IV:

RESULTS

Item and Scale Development

Focus Groups

The data from the focus groups revealed that the construct grit was comprised of similar elements as defined by Duckworth et al. (2007). However, there were many smaller elements that, while not completely separate from the factors described by Duckworth et al., were likely to add more depth to the items in the Grit-V and would enable them to capture more of the true essence of grit. The total time of the transcriptions from the three focus groups was 1 hour, 42 minutes, 21 seconds. The following descriptions highlight the information that I found most useful and noteworthy in defining the various factors of grit that emerged from the focus group data.

When asked what their understanding of grit was, all participants agreed that perseverance of effort was the core factor of grit. Therefore, perseverance was one of the themes that emerged throughout the transcriptions. The following three excerpts from participants highlight this.

Gerald, 22 years old: Someone with grit, to me, is able to keep going regardless... Doesn't matter the hurdle, if they want something, they get it. Plain and simple.

Sam, 20 years old: One of my friends is on the football team, and let me tell you, grit is something the coaches talk about a lot. Sure talent is important, but not everybody got that... You gotta set goals and meet them. Someone with grit will get there no matter what the cost. That's not always a good thing, but to be the best sometimes you sacrifice somethings to get there.

Janet, 27 years old: I wish I had more grit. No matter how hard I try, I just can't keep at it sometimes. And that's the difference... Sometimes you need to keep your eyes on the prize, and with time you will get there.

The feeling of not being able to stay interested and focused from Janet's testimonial led to the development of item 13 from the pilot study. One element of grit is being able to have consistency of interest over longer periods of time, and Janet's situation helped create the situation used in this item.

Most participants described someone with grit as an individual with the mindset of completing their goals even in the face of adversity. Something that may be a nuance to this is that a very "gritty" person may do anything to accomplish their goals, and this includes doing things that may be detrimental to their well-being or may have a negative impact on other aspects of their life. Tabitha, 31 years old, and Ramone, 24 years old, commented:

Tabitha: I feel like I had almost too much grit when I was in college not too long ago. I would force myself to spend hours in the library to get stuff done, even of that meant literally sleepless nights. One semester I lost 10 pounds, but not by diet and exercise mind you. The pressure I put on myself to get through nursing with near perfect marks really put my body to the test.

Ramone: This is more related to sports than study, but still the same. I lost my girlfriend last year because I focused so much time on conditioning for my intramural soccer team. I was so determined to make captain that I spent every spare moment in the gym and on the field. I ended up co-captain, but I think my then-girlfriend had a point that somethings are not worth it. But at the time... I wanted it so badly.

The comment from Ramone led to the initial development of item 6 for the pilot study. While the context is not entirely the same, the item includes a scenario where a student would need to endorse a response about how they would spend their time regarding the completion of a goal.

Another theme that emerged from the data was continual levels of interest. Nearly all of the focus group members mentioned long term goals rather than short term goals. This also matched Duckworth's factor "consistency of interest." While similar to perseverance of effort,

this has more to do with staying focused and interested in something, as well as not getting distracted by other goals, either long or short-term. Of all the instances where consistency of interest was coded in the data ($N = 17$), 15 were connected to long term goals, while the remaining two were related to short term goals. Some of the anecdotal evidence related to consistency of interest related to maintaining a high GPA, keeping a healthy social life with their classmates, and not changing their major, and these examples were worked into various items.

Tenacity was one of the major themes that developed throughout the process. Tenacity was brought up in two of the three focus groups as a solid factor in grit. Therefore, I attempted to capture this in some of the items (e.g., items 2 and 4). However, when asked what the difference between grit and tenacity, most focus group members struggled to offer a concrete distinction.

The following excerpts are some of the attempts to distinguish these two concepts:

Amy, 31 years old: Based on how we are talking about grit, the difference I'm seeing is the obstacles... To be gritty, one needs to overcome obstacles, something in their way. Tenacity perhaps is the same thing minus the obstacles. A tenacious person is determined, but not necessarily facing any adversity.

Demitria, 26 years old: I see tenacious as a positive version of stubbornness. The main difference is a stubborn person won't change even if it's for the better. A tenacious person will adapt. In that way both are strong willed, but the lack of flexibility makes someone stubborn rather than tenacious.

After unpacking the meaning from the anecdotes, I decided that tenacity is a part of perseverance of effort. I believe tenacity is closely enough related to perseverance that it does not warrant being a separate facet of grit. However, this nuance could prove useful in writing items and response as a way to distinguish a tenacious person versus a gritty person.

The last theme that emerged was resilience. Again, after going through the data, I decided that when describing the final facets of grit, resilience falls under perseverance of effort. To distinguish resilience and perseverance, I asked the participants to provide personal examples of

someone they know who is resilient, and I compared my notes with the examples of people with high levels of perseverance. The following excerpts are ones I found particularly useful:

James, 27 years old: My brother and I lost our mom while attending college. I completely fell apart, but my brother managed to keep it together. His outlook on life changed...his goals changed as well. But he did keep a level head. Death affects people differently, but perhaps he was the more resilient one. It still affected him, but not in a negative way as it did me.

Sue, 19 years old: I have a slight learning disability...I'm dyslexic. I was put in special reading and math classes growing up. I still managed to go to the university I wanted. I was constantly faced with doubt, but that didn't keep me from staying positive.

In addition to identifying themes of grit, it was important to gather insight into how grit related to achieving your goals. The purpose was to find out if one can be too gritty. Participants were asked what being too gritty would be, as well as, if being gritty is enough to be successful in school. In two of the focus groups, it was unanimous that grit, while important, is not enough in and of itself. Laura, 27 years old, commented:

Grit is fine and all, but you need some ability to go with it. You can keep at something all you want, but without the talent or ability to do something, how are you supposed to succeed? I may want to be a doctor, but if biology is my worst subject, maybe being gritty will just hurt me in the long run.

Several participants in all three groups mentioned that it's possible to be too gritty. When this came up, they were questioned about what is too gritty? What does it mean to be too gritty? Here are some examples of what participants shared:

Craig, 22 years old: What if you are gritty but you don't adapt to the situation? I immediately think about working out in the gym. Imagine a guy wants to build up muscle but goes about it in all the wrong ways. Maybe he works out every day but never does the right diet. He can do this forever and not see results because he's not going about it in the right way. His ability holds him back.

David, 19 years old: My best friend is a great example of being overly gritty. He puts his study above everything else. Sure, we're in college to learn and plan for our future, but his social life... is nonexistent. I see being too gritty as not being able to balance things. A too gritty of a person may have a one-track mind, and that could be a problem.

The idea of being too gritty is one that was used to describe the interpretation of the Grit-V. This will be discussed later, but in short, someone with a very high grit score may need to adjust their level of grit. In other words, a high level of grit may not necessarily lead to success without potential ramifications in other aspects of their life.

The last step in the focus group process was the follow-up. Participants who agreed to have a follow-up conversation on the phone were contacted. During the conference call, we discussed the items that were I had written and sent them beforehand via email. Three participants mentioned that the item about hard work was harder to understand than the other items, and that the wording needed to be simplified to ensure future test takers would not misinterpret it. A few participants mentioned that while the colloquial nature of some of the items provided authenticity, it was possible that non-native speakers may not understand the idiomatic phrases (e.g., “being burnt out” and “can’t stand someone”). In fact, one of my participants identifies as a native Spanish speaker, and English is her second language. She had difficulty with two items due to the idiomatic language used.

The results from the focus group data reinforced that grit as a non-cognitive construct is multifaceted. Throughout the focus group process, the unpacking of participants ideas that fit into the target audience provided a wealth of information about these facets that the literature alone could not provide. The follow-up process helped ensure that my interpretations of the participants’ personal experiences were accurate and thorough and allowed for member reflections, a common qualitative tool to test for trustworthiness (Tracy, 2010). The process of going through the focus groups has strengthened my belief that the two-factor model was still the best model to consider. While four facets emerged throughout the process, the essence of these facets can be captured in perseverance and consistency of interest, and in the process of writing

the items they can be captured by the subtle nuances of tenacity and resilience to provide a more well-rounded measure of academic grit.

In summary, the themes that emerged from the data were perseverance, long term interest, tenacity, and resilience. There were 23 coded for perseverance, 17 of long-term interest, 10 of tenacity, and 17 of resilience. After reviewing the focus group data, an initial item pool of vignettes was created ($N = 15$) that were inspired from the anecdotal evidence. To stay consistent with the facets of grit that have been established in the literature, eight of the items measured the facet consistency of interest (items 1, 3, 5, 6, 7, 9, 11, and 13) and seven items measured perseverance of effort (items 2, 4, 8, 10, 12, 14, and 15). The themes for each of the items also stemmed from the anecdotal evidence gathered from the focus groups.

Content Validity

An expert panel was used to provide evidence of content validity. The expert panel went through three revisions. A high degree of score reliability was found between the expert reviewers on the 15 vignettes. The average intraclass correlation (ICC) was .822 with a 95% confidence interval from .625 to .932, $F(14,56)=5.625$, $p<.001$. The results from the expert review can be seen in Table 4.1. Based on the review, items 1, 2, 4, 6, 8, 9, 10, 11, 13, and 14 received an average rating below 2; lower values equal higher relevance to the construct. Additionally, the aforementioned items did not have an expert review score above 2. Therefore, these items are considered to have the best content validity of the initial items that were piloted.

Table 4.1

Summary of Expert Panel Review of Vignettes

Vignette	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Average
1	1	1	1	1	1	1.00
6	1	1	2	2	1	1.40
8	1	1	2	2	1	1.40
10	1	2	1	2	1	1.40
13	1	1	2	1	2	1.40
2	2	2	1	2	1	1.60
9	1	2	2	2	1	1.60
14	2	1	2	1	2	1.60
4	2	2	2	1	2	1.80
11	2	1	2	2	2	1.80
5	2	1	1	3	1	2.00
7	1	2	2	2	3	2.00
15	2	2	3	2	2	2.20
3	3	2	3	3	3	2.80
12	4	3	4	2	3	3.20

Pilot Study Results**Partial Credit Rasch Model**

The item calibrations, fit statistics, chi square and reliability of separation statistics were computed. Table 4.2 displays a summary of the results from the Rasch analysis of the Grit-V scale, specifically the calibration of students ($N=20$), items ($N=15$), gender ($N=2$), race ($N=6$), and SES ($N=6$). The data are summarized using average logit-scale calibrations, standard errors, model data fit statistics, and reliability of separation statistics. The results indicated that items were located lower on the logit scale on average ($M = 0.00$, $SD = 1.03$) in comparison to the students ($M = 0.06$, $SD = 0.67$). This suggested that students have a higher probability of scoring

high on the scale. Upon inspection of the infit (inlier-sensitive or information-weighted fit, Linacre, 2002) and outfit (outlier-sensitive fit, Linacre, 2002), the model fit statistics demonstrated an adequate fit to the model; average Infit and Outfit mean square statistics are near 1.00, and the average standardized Infit and Outfit statistics are near the expected value of 0.00 when the data fit the model (Engelhard, 2013; Smith, 2004). Acceptable fit to the model provided support for the interpretation of students, items, gender, race, and SES calibrations on the logit scale as indicators of their locations on the latent variable.

Table 4.2

Summary of Calibrations and Model-data Fit Statistics

		Student (<i>N</i> =20)	Item (<i>N</i> = 15)	Gender (<i>N</i> = 2)	Race (<i>N</i> = 6)	SES (<i>N</i> = 6)
Calibrations						
Measure (Logits)	<i>M</i>	0.06	0.00	0.00	0.00	0.00
	<i>SD</i>	0.67	1.03	0.06	0.07	0.42
Standard Error	<i>M</i>	0.49	0.41	0.16	0.17	0.22
	<i>SD</i>	0.09	0.03	0.04	0.07	0.05
Model-Data Fit						
Infit <i>MSE</i>	<i>M</i>	1.00	0.99	1.02	1.02	1.06
	<i>SD</i>	0.32	0.70	0.05	0.04	0.32
Std. Infit	<i>M</i>	-0.06	-0.29	0.10	0.11	-0.02
	<i>SD</i>	0.95	2.23	0.42	0.29	1.91
Outfit <i>MSE</i>	<i>M</i>	1.04	1.04	1.07	1.04	1.11
	<i>SD</i>	0.42	0.93	0.08	0.01	0.36
Std. Outfit	<i>M</i>	0.06	-0.17	0.46	0.34	0.40
	<i>SD</i>	0.98	2.38	0.52	0.16	1.85
Separation Statistics						
Reliability of Separation		0.81	0.98	0.85	0.61	0.53
Chi-Square		2742.3*	762.5*	22.7*	16.2	11.2

* $p < .05$

The items estimates were examined next and are displayed in Table 4.3. A higher average score equals an easier item, while a lower score equals a more difficult item. The item difficulty locations ($N = 15$) show that item 11 is the most difficult (*Average Score* = 2.58; $\delta = 1.49$; $SE = 0.41$). In contrast, item four was the least difficult item (*Average Score* = 3.70; $\delta = -0.69$; $SE = 0.40$). The values of the item fit statistics indicated an overall adequate fit to the partial credit model, with some minor unexpected responses from individual students for some items. The highest values of fit statistics were observed in item one (Infit $MSE = 1.96$; Std. Infit = 2.96; Outfit $MSE = 2.15$; Std. Outfit = 1.70); the values of the standardized fit statistics suggested slightly more variation than expected related to these items. The lowest values of the fit statistics were observed in item 15 (Infit $MSE = 0.22$; Std. Infit = -2.37; Outfit $MSE = 0.39$; Std. Outfit = -3.87). These values may not be within the range of expected fit values when data fit the model, so this will be considered when deciding on the final items for the Grit-V; the acceptable range commonly used in Rasch measurement is 0.7 – 1.3 for multiple choice non-high stakes testing, Bond & Fox, 2015).

Table 4.3

Summary of Item Estimates

Item	Observed Average	Measure	SE	InfitMS	InfitZ	OutfitMS	OutfitZ
1	3.35	-0.53	0.39	2.96	5.00	3.85	5.99
2	3.25	-0.23	0.39	1.35	1.23	1.18	0.70
3	3.65	-1.58	0.46	0.86	-0.37	0.88	-0.16
4	3.70	-1.80	0.48	0.80	-0.53	0.67	-0.64
5	3.40	-0.69	0.40	1.91	2.81	2.02	2.76
6	2.75	1.33	0.41	1.15	0.55	0.78	-0.45
7	2.75	1.33	0.41	0.79	-0.58	0.51	-1.33
8	3.45	-0.85	0.40	0.50	-2.23	0.57	-1.54
9	3.15	0.07	0.39	0.52	-1.87	0.57	-1.70
10	3.15	0.07	0.39	0.22	-3.87	0.26	-3.72
11	2.58	1.49	0.41	0.61	-1.24	0.59	-0.94
12	3.10	0.22	0.39	0.71	-0.96	0.66	-1.24
13	2.8	1.16	0.40	0.28	-2.88	0.32	-2.37
14	3.35	-0.53	0.39	0.84	-0.56	0.87	-0.39
15	3.00	0.53	0.39	1.36	1.08	1.88	2.41

*Infit scores not within the acceptable range (0.7 – 1.3)

Next, the student estimates were examined using average rating on the raw score scale, the logit-scale measures (θ), SEs, and model-data fit indices. A summary of the student estimates are shown in Table 4.4. The lowest score possible is 15, and the highest possible is 60. Higher total score equals higher level of grit. The values of the student fit statistics (i.e. student fit) indicated an overall adequate fit to the partial credit model, with some minor departures from expectations for individual students. However, these values are within the range of values for Rasch fit statistics that several researchers have shown to be “normal” when data fit the model (Smith, 2004; Wu & Adams, 2013).

Table 4.4

Summary of Student Estimates

Student	Total Score	Observed Average	Measure	SE	InfitMS	InfitZ	OutfitMS	OutfitZ
1	50	3.33	-0.37	0.6	0.79	-0.78	0.73	-0.63
2	43	2.87	-0.54	0.39	0.68	-1.08	0.64	-1.01
3	43	2.87	-0.14	0.43	1.02	0.17	0.99	0.06
4	54	3.60	1.32	0.48	1.20	0.61	1.61	1.1
5	51	3.40	0.82	0.48	1.80	2.06	1.92	2.22
6	50	3.33	-0.37	0.60	0.79	-0.78	0.73	-0.63
7	50	3.33	-0.37	0.60	0.79	-0.78	0.73	-0.63
8	50	3.33	-0.37	0.60	0.79	-0.78	0.73	-0.63
9	50	3.33	-0.37	0.60	0.79	-0.78	0.73	-0.63
10	50	3.33	-0.37	0.60	0.79	-0.78	0.73	-0.63
11	43	2.87	-0.54	0.39	0.68	-1.08	0.64	-1.01
12	44	2.93	-0.03	0.51	1.12	0.45	1.15	0.51
13	42	2.80	-0.94	0.46	0.73	-0.80	0.74	-0.80
14	43	2.87	-0.06	0.43	1.02	0.17	0.99	0.06
15	42	2.80	-0.94	0.36	0.65	-1.11	0.68	-0.95
16	43	2.87	-0.14	0.43	1.02	0.17	0.99	0.06
17	53	3.53	1.29	0.48	1.33	0.93	1.57	1.20
18	52	3.47	1.19	0.36	1.21	0.61	1.69	1.03
19	50	3.33	0.69	0.49	1.58	1.60	1.60	1.65
20	48	3.2	0.07	0.46	1.22	0.75	1.27	0.88

In order to facilitate the interpretation of the gender subgroup locations, the location of the male subgroup was fixed to 0.00 logits. The results for the gender differences are displayed in Table 4.5. The results indicated that, on average, the items were more difficult for males than females. For males, the average logit-scale location was 0.00 logits ($SE = 0.19$); for females, the

average logit-scale location was 0.06 logits ($SE = 0.13$). Overall, the female subgroup was located .08 logits higher than the male subgroup on the logit scale. However, 0.06 logits was a very low difference; therefore I cannot conclude that there is any meaningful difference in grit level between males and females. The model-data fit statistics for the gender subgroups indicated an overall adequate fit to the partial credit Rasch model, with a reliability of separation value equal to 0.84.

Table 4.5

Summary of Gender Differences

Gender	Observed Average	Measure	SE	InfitMS	InfitZ	OutfitMS	OutfitZ
Male	3.04	0.00	0.19	1.05	0.39	1.12	0.82
Female	3.15	0.06	0.13	0.98	-0.20	1.01	0.09

Race and SES were not statistically significant facets in the pilot study. Therefore, I did not further examine the results for race and SES for the pilot study data.

Construct Validity

To show evidence of construct validity, a confirmatory factor analysis was conducted to check if the items fit the two-factor model proposed by Duckworth and Quinn (2009). This model had five perseverance of effort items that loaded on a single factor, and five consistency of effort items that loaded on a second factor. In the model, seven perseverance of effort items were loaded on one factor while eight consistency of effort items were loaded on a second factor. The results showed that the data from the pilot did not adequately fit the model, $\chi^2 = 1393.20$, $df = 77$, $p = .011$, $RSMEA = .011$, $GFI = .60$, $CFI = .61$. However, it is possible that the model did not fit due to the relatively small sample of the pilot study. Table 4.6 contains the item loadings for the two-factor model. Even though the model was not statistically significant, it appeared as

though there were items with relatively strong loadings. This means that with a larger sample, the items may have higher loadings than existed in the pilot, and the model may become significant with the retained items.

Table 4.6

Item Loadings for the Two-Factor Model

Variable	Factor 1	Factor 2	Unique Variance
Item 1	.442		.453
Item 5	.214		.334
Item 6	.546		.543
Item 11	.553		.445
Item 12	.223		.253
Item 13	.415		.298
Item 15	.112		.243
Item 2		.220	.445
Item 3		.465	.212
Item 4		.312	.544
Item 7		.294	.324
Item 8		.320	.554
Item 9		.425	.675
Item 10		.213	.564
Item 14		.314	.467

A principal component analysis was conducted to check the dimensionality of the initial item pool of the Grit-V. Upon inspection of the eigenvalues, the results showed that the initial item pool contained two factors (factors with eigenvalues less than 1.0 were omitted). The extraction sums of the squared loadings for the two-factor solution was 72.4%, while the one-factor solution retained 61%. The factor solution can be seen in Table 4.7. Several of the items that fall into the second factor are ones that the expert panel deemed the least relevant to the

construct. However, caution in interpreting the factor solution is warranted because of the small sample size. Based on these preliminary results, the factor structure of the Grit-V appeared different than what was hypothesized.

Table 4.7

Principal Component Analysis Factor Solution

Variable	Factor 1	Factor 2
Item 1	.566	.221
Item 2	.465	.162
Item 3	.221	.543
Item 4	.465	.321
Item 5	.312	-.043
Item 6	.715	.431
Item 7	.111	.463
Item 8	.487	.230
Item 9	.767	.121
Item 10	.534	-.013
Item 11	.650	.414
Item 12	.222	.563
Item 13	.335	.479
Item 14	.446	.153
Item 15	.164	-.756

Convergent Validity

To find evidence of convergent validity, a correlation study was conducted. The correlations between the Grit-V, Grit-S, NSSE, BFI, and AMS-C were computed. For Grit-V and Grit-S, there was a Pearson correlation coefficient of .430, $p < .001$; there was a significant positive relationship between Grit-V and Grit-S. Therefore, both scales appear to be measuring grit similarly. For Grit-V and NSSE, there was a Pearson correlation coefficient of .518, $p =$

0.01; there was a significant positive relationship between Grit-V and student engagement. Students with high levels of grit are also likely to have high levels of student engagement. For Grit-V and AMS-C, there was a Pearson correlation coefficient of .480, $p = .03$; there was a significant positive relationship between grit and academic motivation. Students with high levels of grit are also likely to have high levels of academic motivation. The relationships between Grit-V and the BFI dimensions (Conscientiousness, Neuroticism, Extraversion, Agreeableness, and Openness to Experience) were not significant. Results are shown in Table 4.8.

Table 4.8

Summary of Correlation Results with Grit-V

Variable (measure)	Pearson Correlation
Grit-S (Grit)	.430*
NSSE (Student Engagement)	.518*
AMS-C (Academic Motivation)	.480*
BFI (Conscientiousness)	.342
BFI (Neuroticism)	-.012
BFI (Extraversion)	.243
BFI (Agreeableness)	-.085
BFI (Openness to Experience)	.231

* $p < .05$

Since there was a significant positive relationship between the Grit-V and Grit-S scales, this suggests that the two scales are measuring the same construct. Additionally, the significant positive relationships between grit, student engagement, and academic motivation provide evident of convergent validity. It is expected that similar results would be found in the final version of the Grit-V. The non-significant relationship between grit and the BFI dimensions were unexpected; this could imply that grit is not directly related to the personality traits. For the final version of the Grit-V, I expect to find significant relationships.

Final Item Revisions and Selection

Once the results of the pilot study were collected and reviewed, the selection process for the final item pool for the Grit-V was completed. The primary basis for item selection was based on the expert content review. Items 1, 2, 4, 6, 8, 9, 10, 11, 13, and 14 from the initial item pool (see Appendix B) had the highest relevancy to the construct, all receiving an average relevancy rating of less than 2 (lower scores equal higher relevancy to the construct). In addition to the review, the results from the Rasch analysis were considered. Items 1 and 15 had the highest fit (more consistent than expected response pattern) and lowest fit (less consistent response pattern than expected) statistics, respectively. Item 1 was retained as it received the highest relevancy to the construct from the expert review. However, item 15 was removed since it received a low relevancy score in addition to having the lowest fit statistic.

Even though the principal component analysis had a small sample, these results were reviewed to facilitate the decision-making process for the final item pool. After consulting with the expert panel about the factor solution, it was decided to check if a one-factor solution fit better and remove any items that had a high factor loading on the second factor. The items with the highest relevancy loaded on factor 1. Therefore, items 3, 7, 12, and 15 were removed from the final item pool. Items 6, 11, and 13 loaded high on both factors. However, since those items had high relevancy scores, they were retained.

The final item pool consisted of ten items (see Appendix C). Items 1, 6, 13, 9, and 11 were designed to measure the consistency of interest facet, and 8, 10, 2, 14, and 4 are designed to measure the perseverance of effort facet. As an addition precaution, a bogus response was created in one of the items. A bogus response is a screening technique, and the purpose is to help ensure that the responses from the students are reliable (DeSimone, Harms, & DeSimone, 2015).

If the students select the bogus response, their results will not be used in the final analysis as their responses to the scale may not be reliable. For an example, the bogus response for item one is *I think keeping a balanced checkbook is more important than staying active and fit*. It has no connection to the vignette, so if a student selects this response it is likely they are not providing appropriate responses.

Final Study

Descriptive Statistics

Prior to running the analyses for the final study, descriptive statistics were collected. Table 4.9 contains a summary of the descriptive statistics for the dependent and independent variables, and Table 4.10 contains the correlation matrix for each of the measures. Overall, the mean scores from the scales were relatively near the midpoint of the minimum and maximum values with the exception of agreeableness. This suggested that on average, the sample population had higher than expected levels of agreeableness. The sample also had a high percentage of students who did not change majors (70%). There was a relatively even distribution of students who did and did not drop classes (45% and 55%, respectively).

Table 4.9

Descriptive Statistics

Variable	Mean	SD	Variance	Range	Min.	Max.
GPA	3.26	.468	.220	1.9	2.300	4.200
Grit-S Total Score	35.75	2.051	4.205	9	32	41
Grit-V Total Score	26.62	3.626	13.148	23	13	36
NSSE Total Score	90.7	11.950	142.805	35	72	107
AMS-S Total Score	90.9	11.949	142.785	40	71	111
Extraversion	26.9	1.138	1.295	6	24	30
Agreeableness	36.4	3.375	11.387	13	26	39
Conscientiousness	36.25	3.814	14.548	13	29	42
Neuroticism	19.1	1.046	1.095	3	18	21
Openness to Experience	33.75	3.652	13.343	13	26	39
Classes Dropped	45% dropped / 55% did not drop					
Changed Majors	30% changed majors / 70% did not change majors					

Table 4.10

Correlation Matrix

	Grit-V	Grit-S	NSSE	AMS-C	Extra.	Agree.	Consc.	Neuro.	Open.
Grit-V	1								
Grit-S	.591*	1							
NSSE	.555*	.415*	1						
AMS-C	.530*	.390*	.552*	1					
Extra.	.354*	.221*	.260*	.335	1				
Agree.	-.410*	-.110	-.329*	-.535*	-.494*	1			
Consc.	.440*	.190*	.415*	.480*	.110	-.410*	1		
Neuro.	-.062	-.121*	-.126	-.084	-.329*	.366*	-.245*	1	
Open.	.285*	.213*	.523*	.510*	.260*	-.419*	.549*	-.453*	1

* $p < .05$ **Research Question 1****Reliability**

To ensure the Grit-V was functioning correctly, I checked for reliability. The reliability analysis provided a Cronbach's alpha of .698. As a second measure of reliability, the split-half method was used. This analysis yielded a Spearman-Brown coefficient of .701. Therefore, with a Cronbach's alpha very close to .7, and a Spearman-Brown coefficient above .7, it was concluded that the Grit-V has an acceptable level of reliability.

Partial Credit Rasch Model Analysis

Prior to analyzing the data, I examined the variance explained by the model to determine if it was acceptable to conduct a Rasch analysis. Thirty four percent of the variance can be

explained with the Rasch model; therefore, we can assume a sufficient level of unidimensionality in the scale and proceed with the interpretations (Reckase, 1979). The calibrations, fit statistics, chi square and reliability of separation statistics were then examined. Table 4.11 displays a summary of the results from the Rasch analysis of the Grit-V scale, specifically the calibration of students ($N=240$), items ($N=10$), gender ($N=2$), race ($N=6$), and SES ($N=6$). The data are summarized using average logit-scale calibrations, standard errors, model data fit statistics, and reliability of separation statistics. The results indicated that items were located lower on the logit scale on average ($M = 0.00$, $SD = 1.03$) in comparison to the students ($M = 0.06$, $SD = 0.67$). This means that participants on average have a higher probability of scoring high on the items. However, as this difference was very slight, the distribution was essentially evenly matched. The model fit statistics demonstrated an adequate fit to the model; average Infit and Outfit mean square statistics are near 1.00, and the average standardized Infit and Outfit statistics are near the expected value of 0.00 when the data fit the model. Acceptable fit to the model provides support for the interpretation of students, items, gender, race, and SES calibrations on the logit scale as indicators of their locations on the latent variable.

Table 4.11

Summary of Calibrations and Model-data Fit Statistics

		Student (<i>N</i> =240)	Item (<i>N</i> = 10)	Gender (<i>N</i> = 2)	Race (<i>N</i> = 6)	SES (<i>N</i> = 6)
Calibrations						
Measure (Logits)	<i>M</i>	-0.42	0.00	0.00	0.00	0.00
	<i>SD</i>	0.56	1.20	0.04	0.29	0.31
Standard Error	<i>M</i>	0.65	0.13	0.06	0.12	0.09
	<i>SD</i>	0.11	0.01	0.01	0.06	0.03
Model-Data Fit						
Infit <i>MSE</i>	<i>M</i>	1.00	1.01	1.03	1.01	1.08
	<i>SD</i>	0.31	0.70	0.05	0.19	0.35
Std. Infit	<i>M</i>	0.04	1.17	0.63	0.14	0.44
	<i>SD</i>	0.88	-0.73	1.17	2.56	2.91
Outfit <i>MSE</i>	<i>M</i>	0.97	0.94	0.97	0.95	1.04
	<i>SD</i>	0.37	0.79	0.06	0.23	0.39
Std. Outfit	<i>M</i>	0.08	-0.21	-0.68	-0.42	0.00
	<i>SD</i>	0.71	1.78	1.16	2.83	5.85
Separation Statistics						
Reliability of Separation		0.78	0.98	0.81	.63	.49
Chi-Square		2042.3*	871.2*	35.1*	46.2	13.4

* $p < .05$

The items estimates and results are displayed in Table 4.12. A higher average score equals an easier item, while a lower score equals a more difficult item. The item difficulty locations ($N = 10$) showed that item 8 was the most difficult (*Average Score* = 2.35; $\delta = 1.69$; $SE = 0.13$). In contrast, item 3 was the least difficult item (*Average Score* = 3.70; $\delta = -2.09$; $SE = 0.15$). It appears as though there was a wide range of difficulty among the items, meaning that the scale should provide a sufficient level of discrimination amongst the students. The values of

the item fit statistics indicated an overall adequate fit to the partial credit model (infit values near 1.00), with some minor unexpected responses from individual students for some items. The highest values of fit statistics were observed in item one (Infit $MSE = 3.18$; Std. Infit = 5.00; Outfit $MSE = 3.03$; Std. Outfit = 9.00); the values of the standardized fit statistics suggested more variation than expected related to this item, meaning that the response patterns from students was not as consistent as it should be. The lowest values of the fit statistics were observed in item 6 (Infit $MSE = 0.35$; Std. Infit = -9.00; Outfit $MSE = 0.43$; Std. Outfit = -8.73). These values are not within the range of expected fit values when data fit the model; the acceptable range commonly used in Rasch measurement is 0.7 – 1.3 for multiple choice non-high stakes testing (Bond & Fox, 2015).

Table 4.12

Summary of Item Estimates

Item	Observed Average	Measure	SE	InfitMS	InfitZ	OutfitMS	OutfitZ
1	3.32	-0.56	0.12	3.18*	5.00	3.03	9.00
2	3.25	-0.29	0.12	1.58	3.55	1.42	4.65
3	3.70	-2.09	0.15	0.73	-2.5	0.62	-3.74
4	2.75	1.49	0.13	0.98	-0.19	0.67	-2.58
5	3.45	-1.02	0.13	0.71	-1.39	0.86	-1.75
6	3.15	0.06	0.12	0.35*	-5.01	0.43	-8.73
7	3.15	0.06	0.12	0.71	-3.4	0.38	-9.00
8	2.35	1.69	0.13	0.74	-3.29	0.73	-1.84
9	2.80	1.30	0.12	0.95	-0.38	0.42	-5.84
10	3.35	-0.65	0.12	1.03	0.35	1.09	1.08

*Infit scores not within the acceptable range (0.7 – 1.3)

Next, the student estimates using average rating on the raw score scale, the logit-scale measures (θ), SEs, and model-data fit indices were examined. The lowest score possible was 10, and the highest possible was 40. A higher total score equals a higher level of grit. Overall, the student fit statistics indicated an adequate fit to the partial credit model, with some minor departures from expectations for individual students.

The highest values of fit statistics were observed in five students (Infit MSE = 1.70; Std. Infit = 1.85; Outfit MSE = 1.73; Std. Outfit = 1.45). Students with overfitting responses implies that they are providing more consistent responses than expected. This could be due to only using a few of the rating scale categories. Three students had the lowest values of fit statistics (Infit MSE = 0.66; Std. Infit = -1.03; Outfit MSE = 0.51; Std. Outfit = -0.76). This suggests that these students are providing less consistent responses. They may choose a response from the rating scale that measures high on grit for one item but select a response that measures low on grit for another. Regardless of these deviations, these values are within the range of values for Rasch fit statistics.

To facilitate the interpretation of the gender subgroup locations, the location of the male subgroup was fixed to 0.00 logits. The results for the gender differences are displayed in Table 4.13. The results indicated that, on average, the items were more difficult for males than females. For males, the average logit-scale location was 0.00 logits ($SE = 0.06$); for females, the average logit-scale location was 0.08 logits ($SE = 0.01$). Overall, the female subgroup was located .08 logits higher than the male subgroup on the logit scale. However, 0.08 logits was a very low difference; therefore, I could not conclude that there was any meaningful difference in grit level between males and females. The model-data fit statistics for the gender subgroups indicated an

overall adequate fit to the partial credit Rasch model, with a reliability of separation value of 0.81.

Table 4.13

Summary of Gender Differences

Gender	Observed Average	Measure	SE	InfitMS	InfitZ	OutfitMS	OutfitZ
Male	3.11	0.00	0.06	1.03	0.63	0.97	-0.68
Female	3.20	0.08	0.01	0.05	1.17	0.06	1.16

The facets race and SES were not statistically significant. Therefore, race and SES were not significant facets in the model.

One of the concerns with the data was item one. With an Infit MSE of 3.18, this item was underfitting; this implies that there was less consistency, and thus more variation, than expected (Bond & Fox, 2015). This suggested there was too much variation in the response pattern, and that there may be some irregular, and perhaps more extreme, response patterns with this item. This also indicated that this item could possibly be distorting or degrading the measurement system. Also, item 6 was overfit, which suggested the response pattern is too predictable, leading to an inefficient item (Wright, 1996).

The residuals of item 1 were inspected graphically due to its high infit statistics. The results are shown in Figure 4.1. The purpose of this graphical representation of the residuals was to show where the disturbances are occurring, rather than simply stating that disturbances exist. The plot contains the observed probabilities across the range of the latent variable; the red line shows the expected score ogive, the blue line displays the empirical ICC, and the black lines show the upper and lower 95% two-sided confidence interval of expected responses. The figure

shows that there are measurement disturbances in this item. Students with theta values (ability levels) between 1 and 3 appear to be the cause of the measurement disturbances.

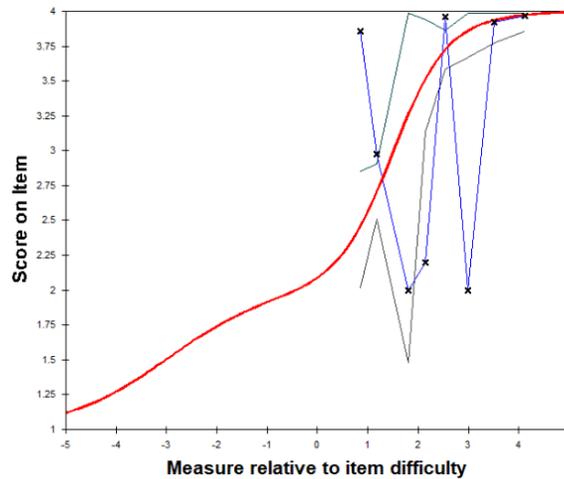


Figure 4.1. *Graphical representation of item 1 residuals*

Additionally, the DIF contrast scores were all less than 1.00 logit units. This suggested that the response probabilities were sufficiently explained by the latent trait (i.e., other variables are not influencing the responses) (Bond & Fox, 2015). The comparisons between groups (e.g., gender, race, and SES) are not problematic. This implies that the scale was assessing grit equally amongst the different groups.

Category Ordering and Precision

Numeric evidence was gathered to provide evidence of rating scale structure. Table 4.14 contains the threshold distances for the items.

Table 4.14

Threshold Distances for Grit-V Items

Item	$ \tau_2 - \tau_1 $	$ \tau_3 - \tau_2 $
1	5.95*	2.9
2	0.47*	0.94
4	1.51	5.45*
8	1.27	2.54

*Threshold absolute values that are not within the 1.4 – 5.0 range

Items 3, 5, 6, 7, 9, and 10 did not have threshold estimates. This is because in this sample the participants did not use all the categories for all of the items. The recommended ranges are between 1.4 and 5.0 (Linacre, 2002). Items 1 and 2 had a higher absolute value for $\tau_2 - \tau_1$ than is recommended. Item 4 had a higher absolute value for $\tau_3 - \tau_2$ than is recommended.

Numeric evidence was gathered to show category precision. Table 4.15 contains the frequencies each category was selected by participants. Categories that are not used, or categories that are dominating a particular rating scale, are of concern.

Table 4.15

Frequencies of Category Observations

Item	Category 1	Category 2	Category 3	Category 4
1	0%	30%	8%	63%
2	0%	20%	35%	45%
3	0%	0%	30%	70%
4	10%	10%	75%	5%
5	0%	0%	55%	45%
6	0%	0%	85%	15%
7	0%	0%	85%	15%
8	0%	40%	50%	10%
9	0%	20%	80%	0%
10	0%	0%	65%	35%

Category 1 was rarely selected across all items. This may not be too alarming, however, as few students in the sample had low levels of grit; the majority had medium and high levels of grit. Therefore, these students would not likely endorse categories that are low on the grit scale.

Numeric and graphical evidence was gathered to show evidence that the rating scale had increasing average threshold values in each category. Overall, the logit-scale locations were increasing across the categories. This show means that the rating scale is functioning as intended. The average logit-scale measures for each category are displayed in in Table 4.16.

Table 4.16

Average Logit-scale Locations Across Items and Categories

Item	Category 1	Category 2	Category 3	Category 4
1	2.53	1.91	1.43	2.30
2		-.36	-.78	.47
3			.23	1.34
4	-.70	-1.04*	-.01	1.91
5			-.60	.23
6			-2.24	-.75
7			-2.12	-1.42
8		-1.88	-.45	.79
9			.58	1.88
10			-.57	-1.01*

*Categories that are not increasing in order as expected

Again, since the majority of students in the sample did not have low grit levels, they did not endorse categories that were low on the rating scale. This was likely the cause of having missing values for the category ratings.

Dimensionality

To check the dimensionality of the Grit-V, as well as find the most parsimonious model, a confirmatory factor analysis was run. The results are shown in Table 4.17. Similar to the pilot study, the results showed that the data did not adequately fit the model, $\chi^2 = 2783.57$, $df = 35$, $p = .019$, $RMSEA = 0.26$, $GFI = 0.64$, and $CFI = 0.62$. Therefore, the factor structure originally proposed by Duckworth and Quinn (2009) was not retained in the Grit-V.

Table 4.17

Grit-V Item Loadings for the Two-factor Model

Variable	Factor 1	Factor 2	Unique Variance
Item 1	.418*		.253
Item 4	.346*		.463
Item 6	.241		.345
Item 8	.273		.295
Item 9	.215*		.314
Item 2		.120*	.376
Item 3		.272	.511
Item 5		.290*	.461
Item 7		.123*	.312
Item 10		.327	.400

* $p < .05$

Since the factor solution was not significant in the Grit-V, I decided to explore the factor solution for the Grit-S. The results are shown in Table 4.18. Interestingly, the results showed that the data from the Grit-S also did not adequately fit the model, $\chi^2 = 1783.57$, $df = 35$, $p = .027$, $RMSEA = 0.29$, $NFI = 0.52$, and $CFI = 0.56$. Since the data from both scales did not sufficiently fit the model, it is possible that the factor structure of the scales is different due to the difference in target populations used in Duckworth and Quinn's research (undergraduate university students versus National Spelling Bee Finalists and West Point Cadets).

Table 4.18

Grit-S Item Loadings for the Two-factor Model

Variable	Factor 1	Factor 2	Unique Variance
Item 1	.221		.342
Item 3	.324*		.396
Item 5	.271		.299
Item 7	.198*		.242
Item 9	.226		.333
Item 2		.135	.317
Item 4		.207*	.441
Item 6		.211*	.401
Item 8		.194*	.303
Item 10		.285	.222

* $p < .05$

Next, I conducted a single-factor CFA to check the factor structure of both the Grit-V and the Grit-S since the two-factor solution did not fit. The results of the CFA for the Grit-V and Grit-S are displayed in Table 4.19.

Table 4.19

Summary of the Single-factor CFA Model for Grit-V and Grit-S

Scale	Variable	Factor
Grit-V	Item 1	.611*
	Item 2	.500*
	Item 3	.512*
	Item 4	.473*
	Item 5	.387*
	Item 6	.511*
	Item 7	.437*
	Item 8	.433*
	Item 9	.572*
	Item 10	.518*
Grit-S	Item 1	.447*
	Item 2	.511*
	Item 3	.407*
	Item 4	.417*
	Item 5	.500*
	Item 6	.476*
	Item 7	.306*
	Item 8	.455*
	Item 9	.337*
	Item 10	.502*

* $p < .05$

For both the Grit-V ($\chi^2 = 161.32$, $df = 35$, $p = .201$, $RMSEA = 0.07$, $GFI = 0.89$, and $CFI = 0.91$) and Grit-S ($\chi^2 = 174.21$, $df = 35$, $p = .111$, $RMSEA = 0.08$, $GFI = 0.86$, and $CFI = 0.88$), the single-factor solution fit the data. Based on these results, it appeared as though grit was a

unidimensional construct when the sample population was undergraduate university students. This suggested that grit may not be a multifaceted construct. In both scales, items that were designed to measure the two facets *perseverance of effort* and *consistency of interest* were measuring the underlying latent trait grit as a single facet rather than as separate facets.

ANOVA

To transpose the Grit-V total scores into a categorical ordinal variable, the scores were recoded into the following ranges: less than 15 = 1 (low); 15 to 24 = 2 (medium); 25 to 34 = 3 (high); and 35 or more = 4 (very high). See Table 4.20 for the grit level ranges.

Table 4.20

Score Ranges for the Grit-V

Points	Grit Level
<15	Low
15 - 24	Medium
25 – 34	High
35+	Very High

Prior to conducting the analysis, the assumption of homogeneity of variances was checked by examining the Levene’s Test of Equality of Variances (Fields, 2013). The Levene Statistic was not significant, $F(2,236) = 60.89, p = .102$; thus homogeneity was assumed. The results from the analysis showed that there were statistically significant mean differences in the model, $F(3,236) = 304.98, p < .001$. The Tukey post hoc tests revealed that there were statistically significant differences between all four levels of grit. Table 4.21 contains the summary of the four grit-level mean estimates from the ANOVA model.

Table 4.21

Summary of Grit-Level Estimates

Grit Level	Mean	Standard Error
Low	13.59	.447
Medium	21.03	.352
High	30.85	.249
Very High	37.01	.159

Completion Time Differences

To determine the average length of time used to complete the Grit-V and Grit-S scales, I checked the time stamps provided by Qualtrics. The average completion time was 6:16 for the Grit-V and 2:34 for the Grit-S. This resulted in the Grit-V taking an average of 3:42 longer to complete than the Grit-S. This was expected as the Grit-S used simple sentence multiple choice items and the Grit-V used longer vignette-based items which take longer to read.

Research Question 2

Cronbach’s Alphas

The results from the internal consistency reliability analysis provided a Cronbach’s alpha of .87 for AMS-C and .82 for the BFI. These are within the acceptable ranges, and therefore I conclude that the scales are functioning as intended.

Comparison of Grit-V and Grit-S on GPA

For comparison one, two simple linear regression analyses were run to test the predictive strength of Grit-V and Grit-S on GPA. The dependent variable was GPA, and the independent variables were Grit-V total scores for Model A and Grit-S total score for Model B. A summary of the results is displayed in Table 4.22. In Model A, there was a significant positive relationship between Grit-V total score and GPA, $b = .813$, $t(238) = 15.752$, $p < .001$ with an adjusted R^2 of

.460. Students with higher grit scores also had a higher GPA. For Model B, the results showed that there was a significant positive relationship between Grit-S total score and GPA, $b = .425$, $t(238) = 7.808$, $p < .001$ with an adjusted R^2 of .201. Again, students with higher grit scores also had higher GPA.

Table 4.22

Summary Results of the Comparison of Grit-V and Grit-S on GPA

Independent	Adj. R^2	Unstd. Beta	SE	Std. Beta	t	Sig.
(Grit-V)	.460	.157	.007	.813	15.752	.000
(Grit-S)	.201	.103	.013	.452	7.808	.000

After examining the results for the regression models, based on the R^2 values, it can be seen that Grit-V was a stronger predictor of GPA than the Grit-S; forty six percent (46%) of the variance was explained in Model A versus 20.1% in Model B. Therefore, the Grit-V was a stronger predictor of academic success in terms of GPA than the Grit-S.

Comparison of Grit-V and Grit-S on Dropped Classes

For comparison two, two binary logistic regression analyses were performed to ascertain the effect of grit on the likelihood that students drop a class. The dependent variable was Dropped Classes (coded 0 for *dropped a class* and 1 for *did not drop a class*), and the independent variables were Grit-V for Model C and Grit-S for Model D. A summary of the results is shown in Table 4.23. Model C was statistically significant, $\chi^2(1) = 77.017$, $p < .001$. The model explained 36.7% (Nagelkerke R^2) of the variance in classes dropped and classified 78.3% of the cases correctly. The Wald criterion demonstrated that grit was a significant predictor. Exp(B) value indicated that when grit increases by one unit, the odds ratio is .576 times as high; therefore, students with higher grit are less likely to drop a class. Model D was also statistically significant, $\chi^2(1) = 8.533$, $p = .004$. The model explained 4.7% (Nagelkerke R^2)

of the variance in classes dropped and classified 35% of the cases correctly. The Wald criterion demonstrated that grit was a significant predictor. Exp(B) value indicated that when grit increases by one unit, the odds ratio is .821 times as high, which indicates that students with higher grit are less likely to drop a class.

Table 4.23

Summary Results of the Comparison between Grit-V and Grit-S on Dropped Classes

Model	Nagelkerke R ²	Beta	SE	Wald	Exp(B)	Predicted % Correct	Sig.
C (Grit-V)	.367	-.552	.074	55.810	.576	78.3	.000
D (Grit-S)	.047	-.197	.071	7.220	.821	35	.000

After comparing the R² values of the two models, the results show that the Grit-V was better at predicting whether or not a student will drop a class than the Grit-S. Additionally, the percentage of cases predicted correctly is higher for Grit-V (78.3% versus 35%). Overall, the Grit-V provided a better predictor of academic success in terms of predicting whether a student will drop a class than the Grit-S.

Comparison of Grit-V and Grit-S on Changed Majors

For comparison three, two binary logistic regression analyses were performed to ascertain the effect of grit on the likelihood that students change majors. The dependent variable was Changed Majors (coded 0 for *changed majors* and 1 for *did not change majors*), and the independent variables were Grit-V total scores for Model E and Grit-S total scores for Model F. A summary of the results is illustrated in Table 4.24. Model E was statistically significant, $\chi^2(1) = 6.726, p = .010$. The model explained 3.9% (Nagelkerke R²) of the variance in classes dropped and classified 72% of the cases correctly. The Wald criterion demonstrated that grit was a significant predictor. Exp(B) value indicated that when grit increases by one unit, the odds ratio

is .857 times as high; therefore, students with higher grit are less likely to change major. Model F was also statistically significant, $\chi^2(1) = 5.943, p = .015$. The model explained 3.5% (Nagelkerke R^2) of the variance in classes dropped and classified 70% of the cases correctly. The Wald criterion demonstrated that grit was a significant predictor. Exp(B) value indicated that when grit increases by one unit, the odds ratio is .831 times as high, indicating that students with higher of grit are less likely to change major.

Table 4.24

Summary Results of the Comparison between Grit-V and Grit-S on Changed Majors

Model	Nagelkerke R^2	Beta	SE	Wald	Exp(B)	Predicted % Correct	Sig.
E (Grit-V)	.039	-.154	.060	6.500	.857	72	.011
F (Grit-S)	.035	-.185	2.861	5.236	.831	70	.022

Based on the R^2 values, the Grit-V was a better predictor at determining the likelihood of a student changing majors than the Grit-S (.039 versus .035). Even though the results are similar, the Grit-V had a slightly higher R^2 and cases correctly classified percentage. However, these differences while statistically significant, may not be practically meaningful.

Comparison between Grit-V, Grit-S, NSSE, AMS-C, BFI Personality Dimensions on GPA

For comparison four, two multiple regression analyses were run to test the predictive strength of Grit-V, Grit-S, NSSE, AMS-C, and the BFI Personality Dimensions scores on GPA; the only difference in the models was Grit-V total scores was as an independent variable in Model G and Grit-S total scores for Model H. The results are displayed in Table 4.25.

For Model G, the model was first run without Grit-V total scores (baseline model). The results from the omnibus test were significant, $F(8,232) = 76.712, p < .001$, with an adjusted R^2 of .561. Then Grit-V total scores were added to the model, and the results were significant,

$F(8,231) = 84.515, p < .001$, with an adjusted R^2 of .717. The independent variables significantly predicted GPA. By adding Grit-V total scores, there was a significant increase in R^2 of .156, $F(1,231) = 7.803, p = .032$.

Since the omnibus results were significant, the individual parameters were examined to compare the strength of the predictors. The strongest predictor in the model was conscientiousness, $t(231) = 5.700, p < .001$. There was a significant positive relationship between conscientiousness and GPA, $b_6 = .195$, indicating those with higher conscientiousness have higher GPAs. The second strongest predictor in the model was Grit-V total scores, with a significant positive relationship between grit and GPA, $t(231) = 4.143, p < .001$. This shows a significant positive relationship between grit and GPA, $b_7 = .167$, indicating that those with higher grit have higher GPAs. The third strongest predictor in the model was AMS-C, $t(231) = 3.903, p < .001$. There was a significant positive relationship between motivation and GPA, $b_2 = .158$. Students with higher academic motivation had higher GPAs. NSSE was also a significant predictor in the model, $t(231) = 3.205, p = .002$. There was a significant positive relationship between student engagement and GPA, $b_3 = .147$, indicating that students with high engagement also had high GPAs. Openness to experience was a significant predictor in the model as well, $t(231) = 1.182, p = .038$, showing a significant positive relationship between openness to experience and GPA, $b_8 = .067$, indicating that students with high openness to experiences also had high GPAs. Agreeableness was the next strongest predictor, $t(231) = -1.022, p = .019$. There was a significant negative relationship between agreeableness and GPA, $b_5 = -.054$. This suggested that students with higher agreeableness have lower GPAs. The last significant predictor was extraversion, $t(231) = .869, p = .046$, meaning a significant positive relationship

exists between extraversion and GPA, $b_4 = .043$, meaning that students with higher extraversion have higher GPAs. Neuroticism was not a significant predictor in the model.

Similar to Model G, Model H was first run without Grit-S total scores (baseline model), then rerun with Grit-S total scores; the model was significant, $F(8,231) = 80.444$, $p < .001$, with an adjusted R^2 of .627. By adding Grit-S total scores, there was a change in R^2 of .066, $F(1,233) = 3.732$, $p = .010$. Grit-S total scores was a significant predictor in the model, $t(231) = 2.613$, $p = .010$. This shows a significant positive relationship between grit and GPA, $b_1 = .139$, indicating that those with higher grit had higher GPAs.

After comparing these models, the Grit-V total scores had a higher unique contribution to the regression model than the Grit-S total scores based on the change in R^2 (.156 versus .066, respectively). Therefore, the Grit-V total scores provided an overall stronger predictor of GPA than the Grit-S total scores.

Table 4.25

*Summary Results of the Comparison between Grit-V, Grit-S, NSSE, AMS-C, BFI Personality**Dimensions on GPA*

Model	Adj. R ²	Unstd. Beta	SE	Std. Beta	t	Sig.
Baseline Model	.561					
AMS-C		.143	.032	.701	3.711	.000
NSSE		.139	.031	1.331	3.004	.011
Extraversion		.045	.041	.111	.891	.027
Agreeableness		-.050	.011	-.822	-.921	.014
Conscientiousness		.144	.021	1.794	5.444	.003
Neuroticism		-.015	.028	-.048	-.602	.481
Openness to Experience		.056	.013	.511	1.002	.022
G	.717					
Grit-V		.167	.012	1.360	4.143	.000
AMS-C		.158	.043	.731	3.903	.000
NSSE		.147	.060	1.487	3.205	.002
Extraversion		.043	.049	.104	.869	.046
Agreeableness		-.054	.022	-.865	-1.022	.019
Conscientiousness		.195	.010	1.994	5.700	.000
Neuroticism		-.025	.038	-.055	-.645	.520
Openness to Experience		.067	.056	.541	1.182	.038
H	.627					
Grit-S		.139	.053	.718	2.613	.010
AMS-C		.151	.036	.792	4.003	.004
NSSE		.140	.044	1.312	3.100	.019
Extraversion		.039	.041	.097	.801	.006
Agreeableness		-.054	.022	-.865	-1.022	.011
Conscientiousness		.171	.021	1.707	5.441	.041
Neuroticism		-.031	.041	-.081	-.837	.322
Openness to Experience		.058	.061	.509	1.097	.000

Comparison between Grit-V, Grit-S, NSSE, AMC-C, BFI Personality Dimensions and Classes Dropped

For comparison five, two binary logistic regression analyses were performed to ascertain the effects of grit, student engagement, academic motivation and personality dimensions on the likelihood that students drop classes (coded 0 for *dropped a class* and 1 for *did not drop a class*). The only difference in the models is Grit-V total scores was an independent variable in Model I, and Grit-S total scores was an independent variable in Model J. The results summary is shown in Table 4.26.

Model I was statistically significant, $\chi^2(3) = 259.046, p < .001$. The model explained 88.3% (Nagelkerke R^2) of the variance in classes dropped and classified 81.3% of the cases correctly. The Wald criteria demonstrated that grit, student engagement, and academic motivation were all significant predictors. For grit, the exp(B) value indicated that when grit increases by one unit, the odds ratio is .016 times as high; therefore, students with higher grit are .016 times less likely to drop classes. For student engagement, the exp(B) value indicated that when student engagement increases by one point, the odds ratio is .072 times as high. This means that students with high levels of student engagement are less likely to drop a class. Lastly, for academic motivation, the exp(B) value indicated that when academic motivation increases by one unit, the odds ratio is .038 times as high, suggesting that students with high academic motivation are less likely to drop a class. The personality dimensions were all non-significant predictors in the model.

For Model J, only Grit-S was a significant predictor. Exp(B) value indicated that when grit increases by one unit, the odds ratio is .513 times as high, and this means that students with higher grit are less likely to drop a class.

Table 4.26

*Summary Results of the Comparison between Grit-V, Grit-S, NSSE, AMC-C, BFI Personality**Dimensions and Classes Dropped*

Model	Nagelkerke R ²	Predicted % Correct	Beta	SE	Wald	Exp(B)	Sig.
I	.883	81.3					
Grit-V			-4.143	1.007	30.352	.016	.000
AMS-C			-2.635	.402	21.989	.072	.000
NSSE			-3.361	.493	26.383	.038	.000
Extraversion			.383	.522	2.853	1.467	.095
Agreeableness			-.517	.107	4.597	.596	.302
Conscientiousness			-.503	.442	3.218	.605	.141
Neuroticism			-.231	.401	1.845	.794	.621
Openness to Experience			-.223	.012	1.767	1.250	.099
J	.352	65.0					
Grit-S			-.667	.107	28.878	.513	.000
AMS-C			-.077	.074	3.295	.926	.069
NSSE			-.047	.044	1.109	.954	.292
Extraversion			.317	.404	14.118	1.373	.340
Agreeableness			-.104	.110	4.113	.901	.076
Conscientiousness			-.113	.041	4.109	.893	.292
Neuroticism			-.265	1.503	1.994	.767	.364
Openness to Experience			-.001	.465	0.015	.999	.081

Similar to comparison two, the Grit-V total scores provided a better predictor for determining if a student will drop a class than the Grit-S total scores. Even though Model J had a strong cases-predicted correct percentage of 65%, Model I had a higher percentage at 81.3%. This suggests that the scores from the Grit-V can better predict if a student is likely to drop a class than the scores from the Grit-S.

Comparison between Grit-V, Grit-S, NSSE, AMC-C, BFI Personality Dimensions and Changed Majors

For the final comparison, two binary logistic regression analyses were performed to ascertain the effects of grit, student engagement, academic motivation, and the personality dimensions on the likelihood that students will change majors (coded 0 for *changed majors* and 1 for *did not change majors*). The difference in the models was Model K had Grit-V total scores as an independent variable, and Model L had Grit-S total scores. The results are shown in Table 4.27.

Model K was statistically significant, $\chi^2(8) = 57.573, p < .001$. The model explained 30.2% (Nagelkerke R^2) of the variance in classes dropped and classified 74.2% of the cases correctly. In this model, academic motivation, agreeableness, and openness to experience were statistically significant predictors. Grit-V was not a significant predictor in the model. For academic motivation, the exp(B) value indicated that when academic motivation was increased by one unit, the odds ratio is .604 times as high, meaning that a student with high academic motivation is less likely to drop a class. For agreeableness, the exp(B) value indicated that when agreeableness increases by one unit, the odds ratio is 1.909 times as high. A student is more likely to change majors if they have high agreeableness. The last significant predictor was openness to experience. The exp(B) value indicated that when openness to experience increases by one unit, the odds ratio is 4.473 times as high. A student with high openness to experience was more likely to change majors.

Model L was not statistically significant, $\chi^2(8) = 17.573, p = .101$. Therefore, none of the independent variables were significant predictors in determining if a student will change majors.

Table 4.27

*Summary Results of the Comparison between Grit-V, Grit-S, NSSE, AMC-C, BFI Personality**Dimensions and Changed Majors*

Model	Nagelkerke R ²	Predicted % Correct	Beta	SE	Wald	Exp(B)	Sig.
K	.302	74.2					
Grit-V			-.468	.583	4.643	.626	.422
AMS-C			-.503	.196	6.567	.604	.010
NSSE			-.455	.211	4.648	.634	.031
Extraversion			-.283	.722	.153	.753	.695
Agreeableness			.647	.302	4.597	1.909	.032
Conscientiousness			-.593	.619	.918	.553	.338
Neuroticism			-.240	.486	.245	.787	.621
Openness to Experience			1.498	.585	6.561	4.473	.010
L	.252	45.0					
Grit-S			-.267	.107	8.878	.766	.120
AMS-C			-.077	.074	3.295	1.080	.069
NSSE			-.047	.044	1.109	1.048	.292
Extraversion			-1.365	1.503	.824	3.916	.364
Agreeableness			1.098	.465	5.575	2.998	.081
Conscientiousness			-2.709	1.731	2.450	15.014	.118
Neuroticism			-1.336	1.662	1.662	3.804	.197
Openness to Experience			-4.334	4.482	4.482	.013	.153

After reviewing the results of the models, both Grit-V and Grit-S total scores were not significant predictors in determining if a student will change majors. Academic motivation was the strongest predictor in the model; an academically motivated student is less likely to change majors. Openness to experience was another strong predictor in the model, suggesting that a student who is open to new experiences is more likely to change their major. Agreeableness was

the next strongest predictor variable; a person with high agreeableness was more likely to change majors.

MANOVA

The results from the MANOVA omnibus test were significant, $\Lambda = .020$, $F(8,236) = 76.761$, $p < .001$, suggesting that at least one of the univariate effects was significant. Therefore, the univariate ANOVA tests were investigated to explore the differences. Table 4.28 contains the results for the MANOVA analysis.

There were significant mean differences for the grit levels on GPA, $F(3,236) = 25.426$, $p < .001$. The Tukey post-hoc results revealed that students with very high grit had significantly higher GPAs ($M = 3.73$) than students with medium and low levels of grit ($M = 3.230$ and $M = 2.842$, respectively). Students with high grit also had significantly higher GPAs ($M = 3.61$) than students with medium and low levels of grit. However, students with high and very high grit did not have GPAs that significantly differed.

There were significant mean differences for the grit levels on student engagement, $F(3,236) = 10.612$, $p < .001$. The Tukey post-hoc analysis showed that students with very high levels of grit had significantly higher levels of student engagement than students with low levels of grit ($M = 98.200$ and $M = 84.510$, respectively). The remaining grit levels did not significantly differ.

There were significant mean differences for the grit levels on academic motivation, $F(3,236) = 5.934$, $p < .001$. The Tukey post-hoc analysis showed that students with high levels of grit had significantly higher levels of academic motivation ($M = 100.750$ and $M = 90.470$, respectively). The remaining grit level means did not significantly differ.

There were significant mean differences for the grit levels on all of the Big Five personality dimensions. The Tukey post-hoc showed that there were differences in means for each of the personality dimensions: extraversion, $F(3,236) = 41.627, p < .001$; agreeableness, $F(3,236) = 13.191, p < .001$; conscientiousness, $F(3,236) = 31.614, p < .001$; neuroticism, $F(3,236) = 22.132, p < .001$; and openness to experience, $F(2,236) = 28.244, p < .001$. When looking at extraversion, students with very high levels of grit had significantly higher levels of extraversion ($M = 28.234$ and $M = 26.200$, respectively). The remaining levels of grit did not significantly differ. For agreeableness, there were significant mean differences between students with very high and low levels of grit ($M = 33.936$ and $M = 37.200$, respectively). However, the remaining levels did not significantly differ. Students with very high levels of grit had significantly lower neuroticism means than students with low levels of grit ($M = 18.780$ and $M = 20.000$, respectively). The remaining levels were not significantly different. Finally, students with very high levels of grit had higher openness of experience means than those with low grit ($M = 36.500$ and $M = 30.192$, respectively). The remaining grit levels did not differ.

Table 4.28

Summary of MANOVA Results

Dependent Variable	Wilks Lambda	F	Df	Sig.	Partial Eta Squared
Omnibus	.031	76.761	24	.000	.706
GPA		25.426	3	.000	.238
NSSE (Student Engagement)		10.612	3	.000	.138
AMS-C (Academic Motivation)		5.934	3	.000	.065
BFI (Agreeableness)		13.191	3	.000	.154
BFI (Extraversion)		41.627	3	.000	.301
BFI (Conscientiousness)		31.614	3	.000	.301
BFI (Neuroticism)		22.132	3	.000	.242
BFI (Openness to Experience)		28.244	3	.000	.297

Summary of RQ2 Results

For research question 2, the results suggested that when trying to predict a student's academic success in terms of their GPA, classes dropped, and changed majors, the Grit-V total scores provided a better predictor than the Grit-S. This conclusion is based off the variance explained in the regression equations (R^2) and beta weights.

However, when the other independent variables were added to the equations, the Grit-V total scores was no longer the strongest predictor of academic success. The multiple regression analysis showed that conscientiousness was the strongest predictor of a student's GPA ($b = .195$), while the Grit-V was the second strongest predictor ($b = .167$).

When trying to classify if a student is likely to drop a class, the Grit-V was the strongest predictor. However, the Grit-V nor Grit-S were significant predictors in classifying if a student is likely to change majors. Instead, academic motivation and openness to experience were the strongest predictors when trying to classify if a student will change majors. The more academically motivated a student is, the less likely they are to drop change majors. Additionally, the more open a student is to new experiences, the more likely they were to change majors.

Lastly, the results from the MANOVA showed that there were significant mean differences in GPA based on the four levels of grit. Students with very high levels of grit had significantly higher GPAs than students from the medium and low groups; students with high levels of grit did not differ significantly from very high levels of grit.

Research Question 3

Criterion Validity

To provide evidence of criterion validity, I tested the predictive strength of grit on academic success in terms of students' GPA, and whether or not they dropped class and changed majors. Based on the results from the regression models in research question 2, I can conclude that there was evidence of predictive validity for the Grit-V. There was a significant positive relationship between Grit-V and GPA; the grittier an individual is, the higher their GPA will be. There was also a negative relationship between Grit-V and classes dropped and changed majors. A person with high levels of grit is less likely to drop classes and change majors than someone with lower levels. Therefore, there is sufficient evidence to suggest the Grit-V has strong evidence of criterion validity. Additionally, based on the R^2 values in the models, the predictive strength of the Grit-V was stronger than the Grit-S (.660 compared to .201). I can conclude that the Grit-V provided stronger evidence of criterion validity than the Grit-S.

Construct Validity

To provide evidence of construct validity, I tested to see if the factor structure of the Grit-V had the same solution as the Grit-S (two-factor model) in previous research. Based on the results from research question 2, both scales did not fit the two-factor model. Rather, the most parsimonious model for both scales was a single-factor solution. The factor solution results also suggested that the Grit-V was providing a better measure of grit than the Grit-S (based on the NFI, CFI, and RSMEA fit indices). Additionally, the factor loadings appeared to be higher overall in the Grit-V than the Grit-S, albeit very similar. On the Grit-V, item five had the lowest loading of .387, being the only item that loaded lower than the acceptable .4 cutoff range (Pett, Lackey, & Sullivan, 2003). The Grit-S had two items that loaded below this threshold (item seven, .306, and item nine, .337).

Convergent Validity

To provide evidence of convergent validity, a correlation study was conducted. The correlations between the Grit-V, Grit-S, NSSE, BFI dimensions, and AMS-C were computed. A summary of the results is shown in Table 4.29. For Grit-V and Grit-S, there was a Pearson correlation coefficient of .591, $p < .001$. Therefore, there was a significant positive relationship between the two scales, which implies they are measuring grit similarly. For Grit-V and NSSE, there was a Pearson correlation coefficient of .555, $p < .001$. There was a strong significant positive relationship between grit and student engagement. For Grit-V and AMS-C, there was a Pearson correlation coefficient of .530, $p < .001$. Thus, there was a significant positive relationship between grit and academic motivation.

In regard to the Grit-V and BFI dimensions, there were significant correlations as well. For Grit-V and Extraversion, there was a Pearson correlation coefficient of .354, $p < .001$; there

was a significant positive relationship between grit and extraversion. For Grit-V and conscientiousness, there was a Pearson correlation coefficient of .440. Thus, there was a significant positive relationship between grit and conscientiousness. For Grit-V and agreeableness, there was a Pearson correlation coefficient of -.410, $p < .001$. Therefore, there was a significant negative relationship between grit and agreeableness. Last, there was a Pearson correlation coefficient of .285, $p < .001$ for Grit-V and Openness to experience. This implies there was a significant positive relationship between grit and openness to experience. The relationship for Grit-V and neuroticism was non-significant, so there was no significant relationship between grit and neuroticism.

Table 4.29

Summary of Correlation Results with Grit-V

Variable (measure)	Pearson Correlation
Grit-S (Grit)	.591*
NSSE (Student Engagement)	.555*
AMS-C (Academic Motivation)	.530*
BFI (Conscientiousness)	.440*
BFI (Neuroticism)	-.062
BFI (Extraversion)	.354*
BFI (Agreeableness)	-.410*
BFI (Openness to Experience)	.285*

* $p < .05$

Next, correlations between Grit-S and NSSE, AMS-C, and the BFI dimensions scores were computed. The results are summarized in Table 4.30. There were significant positive correlations between scores from the Grit-S and NSSE, AMS-C, conscientiousness, extraversion,

and openness to experience, as well as a negative correlation with neuroticism. The relationship with agreeableness was not statistically significant. Overall, Grit-V scores had stronger correlations with the other variables than Grit-S, with the exception of neuroticism.

Table 4.30

Summary of Correlation Results with Grit-S

Variable (measure)	Pearson Correlation
NSSE (Student Engagement)	.415*
AMS-C (Academic Motivation)	.390*
BFI (Conscientiousness)	.190*
BFI (Neuroticism)	-.121*
BFI (Extraversion)	.221*
BFI (Agreeableness)	-.110
BFI (Openness to Experience)	.213*

* $p < .05$

The positive correlations between the scores from the Grit-V, NSSE, and AMS-C are consistent with previous literature (Von Culin, Tsukayama, and Duckworth, 2014). The relationships between Grit-V scores and the BFI personality dimensions scores are also consistent with the literature (Duckworth & Quinn, 2009) except for Neuroticism. Duckworth and Quinn reported a significant negative correlation with grit and Neuroticism. While the relationship between the Grit-V and neuroticism was negative, it was non-significant. Based on these results, there is sufficient evidence not only for convergent validity, but also that the Grit-V has stronger convergent evidence than the Grit-S.

Content Validity

The inter-rater reliability results from the pilot study showed that the expert panel had a high level of agreement with an intraclass correlation of .822. This is evidence that the Grit-V has content validity. However, I do not attempt to provide content validity evidence for the Grit-S, so I am not able to claim that the Grit-V has more evidence than the Grit-S.

Summary of RQ3 Results

The results from research question two suggested that Grit-V not only provided certain evidence of validity, but also that Grit-V provided more validity evidence than the Grit-S. While both the Grit-V and Grit-S provided evidence for criterion validity, the Grit-V provided stronger correlations between grit and academic success than the Grit-S based on the R^2 values. Interestingly, both scales failed to provide construct validity based on the existing factor structure. However, since dimensionality of a construct can change based on the target population, this may not be evidence that both scales failed to provide construct validity. For convergent validity, both scales provided evidence; however, the Grit-V showed stronger correlation between grit and the independent variables, and thus showed stronger evidence of convergent validity than Grit-S. Lastly, through the expert panel review process, evidence of content validity was present as the inter-rater reliability coefficient was very high. However, I was not able to determine if the evidence of content validity in the Grit-V is stronger than in the Grit-S.

Research Question 4

To investigate whether SES is a significant moderator variable in the models, three interaction terms were created by multiplying the SES variables by Grit-V Total Score (Parents'

Education Level, Parents' Occupational Status, and Parents' Annual Income). In all three models, SES was not a significant moderator variable. The results are shown in Table 4.31.

Table 4.31

Summary Results of Moderator Analysis

Model	Adj. R ²	DF	F	Partial Eta Squared	Mean Square	Sig.
M	.700	6	53.842	.505	6.186	.008
Grit-V		1	6.980	.213	.196	.036
Parents' Educational Level		2	9.778	.031	.239	.061
Parents' Educational Level x Grit		2	3.952	.033	.261	.706
N	.807	33	31.310	.834	1.326	.000
Grit-V		1	5.049	.185	.341	.038
Parents' Income		10	12.443	.209	.230	.091
Parents' Income x Grit-V		10	15.720	.217	.242	.089
O	.502	6	41.324	.465	2.321	.002
Grit-V		1	5.223	.190	.254	.000
Parents' Occupational Status		2	4.546	.086	.135	.241
Parents' Occupational Status x Grit		2	3.546	.024	.223	.081

As an additional way to inspect the mean differences in grit score for the different SES groups, an ANOVA analysis was conducted; GPA was the dependent variable and Parents' Educational Level, Occupational Status, and Annual Income were the independent variables. The omnibus test results showed that the model was significant, $F(2) = 66.234, p < .001$. The main effect Annual Income was not statistically significant, $F(2) = 3.098, p = .080$, nor was Occupational Status, $F(2) = 2.988, p = .085$. However, Education Level was statistically significant, $F(2) = 84.155, p < .001$. Students whose parents' highest level of education was high school had the lowest mean grit score of 27.412. In contrast, students who had parents with a graduate level education had the highest level of grit with a mean score of 34.659. There were

statistically significant mean differences between all of the groups besides AA and Undergraduate. The estimates for SES Educational Level is displayed in Table 4.32.

Table 4.32

Summary of Education Level Estimates for Grit

Education Level	Mean	Standard Error
High School	27.412	.200
Some College	30.413	.295
AA	32.587	.301
Undergraduate	31.687	.249
Graduate	34.659	.463

While SES was not a statistically significant moderator variable, there were meaningful differences between the education level groups and grit. The results suggested that a parents' education level had a significant effect on students' level of grit. On average, the more education parents had, the more grit their children had.

CHAPTER V:

DISCUSSION

Item Development Process

Focus Groups

One of the purposes of the study was to use a variety of scale development methods to create valid and reliable score measures of grit. The use of focus groups yielded interesting results. In one respect, the results were consistent with the current literature. The themes that emerged during the focus group sessions (tenacity, perseverance of effort, consistency of interest, and resilience) were similar to the ones Duckworth and Quinn (2009) found. Even though the final decision was to reduce the four themes into two overarching facets, the anecdotal evidence from the focus groups provided much needed insight on the item development process. For example, the scenarios that were used in the items were ones that actual college students experienced. Using the items from the Grit-S as a template, I was able to weave the scenarios from the focus groups in order to create authentic vignettes. I believe this form of item development is not only useful in developing a form of grit measurement, but also can be used in other constructs such as self-efficacy, optimism, resilience, and so forth. Since the focus group participants were from the target population of the Grit-V, the experiences they shared helped me develop items that would better relate to the target population. However, the facets I decided on using in the scale ended up being not being significant factors; rather, a single factor solution was the most parsimonious model. This will be discussed more in the psychometric properties section.

Psychometric Properties

Reliability. The results from the reliability analysis showed that the Grit-V is functioning correctly. While the Cronbach alpha was lower than the expected .7 threshold, the Spearman-Brown coefficient was above this threshold. There is no concrete agreement among researchers on what exactly Cronbach alpha values should be. Taber (2018) provided a list of qualitative descriptors for Cronbach alpha interpretations used by a large number of different scholars; he stated that “the terminology seems to be somewhat arbitrary” (p. 1278). Since the qualitative descriptors in Taber’s meta-analysis for alpha levels between .6 and .7 in leading science education journals include *acceptable*, *satisfactory*, *sufficient*, and *adequate*, the value is close enough to conclude that the scale is working as intended and has an acceptable level of response reliability.

Partial credit Rasch model. The results for the item estimates suggested that there was an adequate spread of item difficulty in the Grit-V. This means that the items were offering an adequate level of discrimination between the different levels of grit, suggesting that the scale should be able to provide a more precise form of measurement of grit. The vignettes were able to tease out the subtle differences in grit levels between students because of the content-specific nature of the items.

Model data fit. While the data fit the Partial Credit Rasch Model overall, there were a few problematic items. Item 1 was underfit, suggesting that there are less consistent responses than expected. However, this may not actually be a problem per se. This item contained the bogus response, reducing the appropriate response pool to three rather than four. This is possibly the cause of the underfit. Since the purpose of the bogus response was to increase participant response reliability (i.e., remove participants who are not providing reliable responses to the

items), removing the bogus response to reduce the underfit does not outweigh the benefit.

Another possible reason could be word choice, as poorly written items, or items that use stronger verbiage, could be a source of unwanted error variance (Worthington & Whittaker, 2006). Item 6 was overfit, meaning that there were more consistent responses than expected. After reviewing the item, there may be a few reasons why there were consistent responses. It is possible that some of the participants had an issue with stronger words like *obsessed* and *abandoned*, and therefore avoided those responses more than others. There is always a risk involved when deciding on the best wording of items, and this has been duly noted in the literature (e.g., DeVellis, 2016; Worthington & Whittaker, 2006).

Facets. Gender was a significant facet in the model with females having a slightly higher level of grit than males on average. Previous research (e.g., Duckworth & Quinn, 2009) found that there were no statistical differences between males and females. However, we should be hesitant to postulate that any meaningful difference between the gender groups existed based on my results. A difference of only .08 on a logit scale is likely not going to have any meaningful impact on the interpretation of students' grit scores. Based on the average scores for males and females, the difference is only one point. This difference is not substantive enough to make claims about grit and gender, and further research would be needed to evaluate grit differences between males and females.

SES and race were not significant facets in the model. This suggested that a student's SES and race were not affecting the level of grit a student had. Duckworth (2016) believed that SES does have an impact on grit. More specifically, she advocated that people from low SES backgrounds may in fact have higher grit because of their upbringing (i.e., having to endure hardships). The findings from this study are indicating a different result. The topic of the

relationship between SES and grit was brought up in the focus groups, but none of the participants thought that SES had much impact. Rather, the focus group members believed that grit is more likely instilled in children by the parents than their SES background. The data from this research also showed that both SES and race were not statistically significant facets.

DIF. The results from the DIF analysis suggested that there were no other variables influencing the responses in a statistically significant way. The comparisons between gender, race, and SES groups were not problematic. This means that the Grit-V is measuring the construct grit equally across those groups. This is important because it fulfills one of the requirements for invariant measurement. If there was a lack of invariance, this would suggest that the Grit-V was not functioning the same way across these groups. Since the DIF results showed that the Grit-V is measuring grit equally across these groups, we have evidence of invariant measurement (Engelhard, 2013). This is especially important when interpreting the results of the Grit-V. Since there is no statistical evidence of DIF, we do not need to consider item bias (differences in group performance due to characteristics irrelevant to the Grit-V) nor item impact (differences in group performance due to characteristics relevant to the Grit-V).

Category ordering and precision. Even though the results from these analyses were not optimal, they may not be too problematic. Regarding the threshold distances, Linacre (2002) suggested that the recommended distances are between 1.4 and 5.0, and there were three threshold distances outside of this range. This was likely due to the participants not endorsing all of the categories from all of the items. Future research with a larger sample population could possibly remedy this, as these threshold distances may then fall within the recommended range.

Concerning the category precision, there seemed to be very few responses in category 1 (low grit level) among the students. This may not actually be problematic however as the

observed average for males and females was above 3. This particular sample had overall high levels of grit, so it would be fitting for them to not endorse item responses corresponding to low grit levels. Additional research with a larger sample may provide different results. Also, the evidence that the rating scale was functioning as intended was promising. Only two items had categories where average logit-scale locations that were not increasing (items 4 and 10).

While these results may not be ideal, they do provide a distinction between the Grit-V and Grit-S. At this point, there is no evidence that the rating scale of the Grit-S is functioning as intended, nor do we know if the threshold estimates are within the recommended ranges. Further research would be needed to both explore these psychometric properties of the Grit-S, as well as investigate if a larger sample of students can provide a more complete picture of the Grit-V's category precision.

Validity

Overall, the Grit-V provided adequate evidence of validity. In this section, different types of validity evidence that exist within the Grit-V are described, as well as discussion of some possible issues and improvements.

Criterion validity. The results showed that grit was predictive of academic success. The Grit-V provided more predictive strength than the Grit-S. Therefore, by using vignettes to enrich the context of the items, I was able to improve upon the existing measure of grit in terms of its ability predict students' success. There is ample research that supports the notion that grit has a positive relationship with success in general (e.g., Duckworth et al., 2007; Duckworth & Quinn, 2009; Duckworth & Gross, 2014), though little previous research has studied the direct relationships between grit and academic success specifically in college students.

Construct validity. The results from the CFA revealed that the higher order bifactor model did not hold in the target population for either the Grit-V or Grit-S. However, this was necessarily a lack of construct validity. There were a few possible explanations as to why the factor structure was different for both scales. First, the fact that both scales did not fit the model is promising. This means that it was not the dimensionality of the Grit-V that was the problem. Since both scales did not fit, it was possible that grit was a unidimensional construct when the target population was undergraduate college students. Further research on grit and college students would be needed to confirm or contradict that grit is unidimensional. Therefore, at this point I conclude that since both scales were unidimensional, the Grit-V has evidence of construct validity.

Convergent validity. For the most part, the results from the correlation studies were as expected. Since the correlation between the Grit-V and Grit-S was statistically significant, I can conclude that the two scales are measuring grit in a similar manner.

The relationship between grit and the personality dimensions were nearly as expected. Four of the five dimensions were in line with the literature (Duckworth & Quinn, 2009), with neuroticism being the exception. The authors noted that neuroticism and grit were negatively correlated; while results in this study showed a negative correlation, it was not statistically significant. Theoretically, it makes sense for these two constructs to be negatively related. A person with high levels of neuroticism often experiences feelings of fear, worry, anger, and depression (Thompson, 2008). Stressors can often be more likely to turn small obstacles into hopeless frustrations for someone with high neuroticism. Since a gritty person can often bounce back from obstacles easily, it would be logical to conclude that a person with high levels of grit would have low levels of neuroticism. The fact that the relationship was in a negative direction is

promising because even though the relationship between these constructs was not statistically significant, it is possible that if used again with a different sample the correlation coefficient would be significant.

Content validity. Grit-V was successful in yielding useful items was because of the use of focus groups and expert panels to validate the content of the items. The study results would encourage future scale developers to strongly consider using expert panel reviews throughout the item development process. In the early stages of item development, the developer is trying to design meaningful items to measure a given construct. The use of focus groups allows the developer to gather authentic materials (e.g., anecdotal evidence) from the focus group members. Once an item pool has been written, the expert panel can provide different perspectives that the developer may not have considered. The process of item revision may involve several iterations, but if the panel is involved throughout this process, the developer is likely going to end up with a better item to construct result. It is important to ascertain inter-rater reliability. The use of an expert panel is predicated on the fact that the experts can reliably rate items similarly. If not, it is harder to compare the quality of the items based on the raters' reviews.

Completion Times for the Grit-V and Grit-S

Qualtrics revealed that the Grit-V took longer to complete on average than the Grit-S (over twice as long, in fact). However, the expert review panel and I did not view this as a drawback. These differences in completion times were discussed, and it was concluded that the extra time to complete the survey is not a drawback. The additional validity gained from using the vignettes outweighed the extra time needed to provide responses to the items.

The Predictive Ability of Grit and the Other Constructs on Academic Success

In this section, the predictive strength of grit and the academic success variables, as well as how grit compares to other constructs in terms of predicting a students' potential academic success are discussed.

Both the Grit-V and Grit-S total scores were successful at predicting GPA (students with high grit tend to have high GPAs). This aligns with the theoretical framework that people with high grit tend to be more successful than those with low grit (Duckworth et al., 2007).

Concerning predictive strength, the scores from the Grit-V was more predictive of GPA than the Grit-S. This is expected because of the use of vignettes, which provided more context rich items than simple sentence item endorsement, thus allowing for a more valid measure of the construct.

However, when the other constructs were added to the model, grit was not the strongest predictor of GPA. Conscientiousness was the strongest predictor of GPA. Therefore, when trying to predict GPA, conscientiousness was more predictive than grit when using the Grit-V. This suggested that a person who is conscientious (i.e., someone who is careful and organized) is more likely to be successful in academics than someone who is easy-going and disorganized (Barrick & Mount, 1991). A conscientious person is typically hard-working and goal-oriented, and thus more likely to achieve their goals (Barrick & Mount, 1991).

Interestingly, openness to experience was the significant predictor of GPA. Openness to experience is typically characterized by a number of different facets, including attentiveness to inner feelings, active imagination, having a preference for variety, aesthetic sensitivity, and intellectual curiosity (Costa & McCrae, 1992; McCrae, 1987). Of these facets, intellectual curiosity offered the best explanation for this relationship. A person who is intellectually curious

seems likely to come up with innovative ways to solve a problem or obstacle, a key feature in a person with high grit.

Regarding the predictive ability on the likelihood of dropping a class, both the Grit-V and Grit-S were significant predictors, with the Grit-V having more predictive strength. This is logical because someone who has high grit is less likely to give up on a class because it is too difficult. Other variables that were strong predictors in the model were academic motivation and student engagement. Again, this is quite logical because students who are motivated and engaged are likely to push through difficult coursework (Vallerand, Pelletier, Blais, Briere, Senecal, & Vallieres, 1992), and this was confirmed in the current study.

Concerning changing majors, grit was not a significant predictor when measured by either grit scale. Rather, the results suggested that academic motivation, openness to experience and agreeableness are the strongest predictors in classifying if a student is likely to change majors. The more academically motivated a student is, the less likely they are to change majors. This makes sense since because a student who is highly motivated will usually give up on a major; they would use their motivation as fuel to continue their studies (Allen, 1999).

It is also logical that a student who has high openness to experience would be more likely to change majors. Someone with high openness to experience typically exhibits signs of intellectual curiosity (Costa & McCrea, 1992), and therefore has a higher chance to be willing to try new things. In the context of academia, openness to experience would likely show itself in the willingness to take a wide variety of coursework, as well as being open to change majors.

Agreeableness had a negative relationship with changing majors. Since agreeableness is related to a person's competitiveness (higher scores equal less competitive and confrontational, (Barrick & Mount, 1991) this would suggest that persons with low agreeableness are more likely

to change majors. Findings support the idea that when a person has a competitive nature, they are less willing to “fail” at something. It is possible that changing majors could be viewed as “failing” to a competitively-minded person, and thus is less willing to change majors and more likely to find ways to complete the program they started.

It is important to add that changing majors is not necessarily a sign of poor academic achievement. In fact, it could be a sign that the individual recognizes their strengths (and weaknesses) and adjusts their studies accordingly. In this sense, being too gritty may actually be harmful to one’s success. Consider a student who starts his college career pursuing economics. Once he gets to his second year, he realizes that math and finance are not his strengths, but rather has shown promise in anthropology. If he were to stick with it simply because he cannot stand to give up and fail, he may in fact never succeed. Or, he may sacrifice too much to finish his program. Rather, if he were to be open to new experiences, he may be more willing to explore different academic avenues and discover something else that suits his skills and talents. Of the three variables used to measure academic success (GPA, classes dropped and changed majors), GPA was the most meaningful. GPA is an overall numeric representation of how well a student has performed in their program. While GPA is not the entire picture of success, it is a big part of what is used in making admission decisions into programs (as well as standardized testing).

The Relationships Between Grit and the Other Constructs

Grit and Student Engagement

The positive correlation between grit and student engagement found in this study was consistent with the literature (Von Culin, Tsukayama, & Duckworth, 2014). In fact, the correlation between grit and student engagement was nearly as strong as the correlation between

the Grit-V and Grit-S. This suggested that there was significant overlap between these constructs. Student engagement contains elements of interest, curiosity, and even passion, so it is not unexpected that strong connections between engagement and grit exist. Grit, however, adds the element of perseverance and overcoming obstacles.

Grit and Academic Motivation

The positive correlation between grit and academic motivation was also very strong. A student with high academic motivation increased initiation of activities, and therefore are more likely to do things that lead to improved performance (Ormrod, 2003). While future researchers can investigate how the different types of motivation (intrinsic, extrinsic, and amotivation) interact with grit, this study offers insights on why these constructs are positively related. A gritty person is likely to be both intrinsically and extrinsically motivated. No statistical evidence has been gathered at this point to show which form of motivation is more related to grit, but on a theoretical level, both make sense. A gritty student would likely be extrinsically motivated to succeed. For example, a gritty student wants to finish their degree program by any means necessary so they can get the job they desire. Intrinsic motivation comes into play when we talk about passion, and passion is a large part of grit (more specifically, passion for long-term goals (Duckworth et al., 2007).

Grit and Conscientiousness

There were strong correlations between grit and conscientiousness in both grit scales. Both constructs share significant similarities. Both gritty and conscientious people are hard-working and goal driven. People with high levels of grit and conscientiousness are also thorough in their work habits, and are persistent in their overall behavior (Duckworth, Weir, Tsukayama, & Kwok, 2012; Ivcevic & Brackett, 2014). MacCann, Duckworth, and Roberts (2009) even

identified perseverance as a factor of conscientiousness through CFA and EFA analyses, which also is also a significant facet in grit. Grit and conscientiousness work in tandem, and it appears to be difficult to mention grit in research without touching upon conscientiousness.

It is possible to be too conscientious, just as it is possible to be too gritty, and recent research supports this claim. Carter, Guan, Maples, Williamson, and Miller (2015) discussed the pitfalls of being too conscientious. They claimed that individuals with very high levels of conscientiousness are linked to maladaptive obsessive-compulsive traits, which lead to negative impacts on their well-being. Costa and McCrea (1992) suggested that too much conscientiousness could lead to excessive workaholic tendencies. In the present study, the findings uncovered anecdotal evidence during the focus groups to support this claim. Many group members provided examples of people they know who perhaps sacrificed too much to accomplish their goals. Further statistical evidence would be needed to make more concrete claims that extreme levels of grit can negatively impact other areas of one's life.

Grit and Neuroticism

Grit and neuroticism had a negative correlation. As noted earlier, it makes sense for these constructs to be negative related. Someone with high grit should have low levels of neuroticism because they do not let small obstacles overwhelm and discourage them. A gritty person is able to deal with stress and obstacles in a productive way and will not let such things impede their long-term goals (Duckworth et al., 2007). A person with high neuroticism, on the other hand, are more vulnerable to stress and tend to experience more negative life events than those with lower levels (Jeronimus, Ormel, Aleman, Penninx, & Riese, 2013). Small obstacles can appear insurmountable, and thus prevents a person from accomplishing their goals.

Grit and Extraversion

Grit and extraversion showed a positive correlation. A gritty person will do what it takes to accomplish tasks. Being enthusiastic and engaged in your coursework can often lead to completion (Zepke & Leach, 2010), and these are traits that an extraverted person is likely to have. Previous research has shown that people with higher levels of extraversion had positive relationships with career satisfaction and log salary (Seibert & Kraimer, 2001). Perhaps future research could investigate if grit was also a factor in similar outcomes.

Grit and Agreeableness

The negative correlation between grit and agreeableness was not as easily interpreted. In Duckworth and Quinn's (2009) work on the development of Grit-S, they found positive correlations, but both Grit-V and Grit-S showed negative correlations with agreeableness in undergraduate students. People with high agreeableness scores typically are kind, considerate, and cooperative (Graziano & Eisenberg, 1997). On the other end of the spectrum, people with low agreeableness are more likely to compete, rather than cooperate, with others. It is this where we begin to see why grit and agreeableness are negatively correlated. Since a gritty student is more likely to do whatever it takes to accomplish their goals, it is sensible to believe that this student would be willing to compete with, or even manipulate, fellow students in order to complete a given task. Findings in this study based on the focus group data, indicated that it was more likely to occur in students with high and very high levels of grit.

Grit and Openness to Experience

The relationship between grit and openness to experience was weak to moderate. It is important to note that openness to experience may be multidimensional (Griffin & Hesketh, 2004). The authors found that while openness to external experience can positively affect one's

work productivity, openness to internal experience may have negative effects. Further investigations are needed to see how these two facets of openness of experience function in academic settings before claims on the relationship between grit and openness to experience can be made. However, a student who is open to try new things may have a high chance to accomplish their goals, rather than someone who is closeminded. Often times thinking outside the box is the key in finding a solution to a problem, and someone with high openness to experience may have an easier time finding that solution.

Grit-V Guide

The following section is a description of how the Grit-V should be used and how scores should be interpreted. There are three primary uses for the Grit-V: potential supplemental admission criteria, counseling assistance for course planning, and future grit research. In this section, how Grit-V can be used as an additional tool to assist in admission decisions, as well as, how guidance counselors can use this as a tool to help better guide students on the right academic track are discussed. Researchers can also use the Grit-V in their future grit research studies.

Potential Admissions Criteria

The use of non-cognitive skills in college admissions is not an entirely new concept. For example, academic institutions such as the University of Norte Dame are beginning to use ETS's Personal Potential Index to gather a more complete picture a student's potential in a graduate program. The Grit-V is different from the Personality Potential Index in that it focuses on one specific skill rather than multiple skills. However, the Grit-V is not meant to be used in isolation; rather, it should be used as supplemental material during the process of making admission decisions.

It is also possible to use the Grit-V to compare potential scholarship candidates. When deciding on award scholarships, be it athletic, academic, and so forth, it is not always a clear choice on who are the best candidates. The Grit-V could be used to assess which candidate is more likely to succeed in their college career. Admissions offices and scholarship awarders hope that the students they select will succeed, and the Grit-V could be an additional measure of the students' potential.

Counseling

It is common for students to seek help from guidance counselors when planning their program of study. Counselors could use the Grit-V as a tool to guide their advice and support given to students. The results from this study showed that students with high grit were less likely to drop classes. Therefore, the Grit-V could also potentially be used as a diagnostic tool in college counseling. Let us consider students who have strong GPA and test scores but has low grit. These students could be flagged as having a possible risk of struggling if they had to deal with life and academic obstacles throughout their academic careers. Guidance counselors can provide these students support by helping them improve their non-cognitive skills like grit. In terms of helping students grow their grit, Perkins-Gough (2013) interviews Duckworth and provides some suggestions, such as creating a desire to constantly challenge oneself, always find a sense of purpose in what you do and maintain hope. These are the types of things guidance counselors and teachers alike can instill in their students to develop stronger non-cognitive skills.

Future Grit Research

Researchers interested in investigating grit in college students could benefit from using the Grit-V rather than other forms of grit measurement. Since the Grit-V is specifically designed to assess grit in undergraduate college students, it offers a more valid measure of grit than the

Grit-S, which is intended for a more general target population. Grit has become a popular topic in education in recent years, so offering a form of measurement that is tailored to the education field may help strengthen research findings with this target population.

If future researchers become interested in grit in other domains, such as business, sports, or social behavior, additional scales could be developed to help measure grit in that specific context. Using the Grit-V as an example, researchers could write items that are content-related to a given area of interest. While the Grit-S is a quick way to measure individuals' grit, a vignette-based scale may provide more information to researchers in those content areas.

Interpretations

The Grit-V scores ranged from 10 to 39 points (4 points per item, with item 1 having 3 points due to the bogus response). Table 5.1 shows the score breakdown for the various levels of grit. This description of each level of grit is given to better understand Grit-V score interpretation. See Appendix C for the scoring guide.

Table 5.1

Score Point Breakdown for the Grit-V

Points	Grit Level
<14	Low
15 – 24	Medium
25 – 34	High
35+	Very High

Low. Students who score less than 14 points have very low levels of grit. Students with scores in this range are typically selecting item responses from mainly the first level of grit, with some responses from the second level (1s and 2s). These students would be less likely to perform well in their undergraduate program when facing difficult obstacles. This would most likely have

adverse effects on their academic success, including lower GPA and having a higher likelihood to drop classes due to course difficulty. Students with very low grit are very likely to give up on difficult tasks and allow obstacles to discourage them from accomplishing their long-term goals.

Medium. Students who score between 15 – 24 points have medium levels of grit. Students with scores in this range are typically selecting item responses mainly from the second level of grit, with some responses from the third level (2s and 3s). Students with medium levels of grit are likely to overcome obstacles when presented. While students with medium scores still may drop a class due to difficulty, they are less likely to do so than low and very low grit students. Additionally, students with medium levels of grit are less likely to let distractions affect their long-term goals. It is also easier for these students to stay interested in the activities they pursue.

High. Students who score between 25 – 34 points have high levels of grit. Students with scores in this range are typically selecting item responses mainly from the third level of grit, with some responses from the fourth level (3s and 4s). Students with high grit are not likely to drop classes due to difficulty, and they are significantly less likely to let distractions get in the way of their long-term goals. They will stay focused on the tasks and goals they pursue more than those with medium and low levels of grit.

Very high. A student who scores 35 or more points has very high levels of grit. Students with scores in this range are typically selecting mostly item responses from the fourth level of grit. These students have the lowest likelihood to allow distractions to interfere with their long-term goals and drop classes due to difficulty. It is very unlikely that students with very high grit will encounter obstacles too large to overcome throughout their studies. Their interests are likely to be consistent over time and not waver when new potential opportunities present themselves.

In addition to the positive effects of having high levels of grit, there can be too much grit. In the very high range, it is possible that too much grit can lead to over-grit. Over-grit is when grit becomes personality traits related to stubbornness and a lack of self-regulation, or control of one's self. Specifically, being over-gritty could cause an individual to have reduced behavioral self-regulation, or the ability to act in one's best interest; in other words, it's the ability to feel one particular way, but act in another (Stonsy, 2011). An over-gritty student may continue to go down a career path even though it is in their best interest to stop and pursue a different academic track. While being persistent and putting in maximum effort is typically a positive trait, it is also important to know when it is better to cut your losses to move to something else. The relationship between grit and self-regulation in this study was not the focus, but future researchers are encouraged to investigate this relationship further.

Making admission decisions. Higher grit scores were connected to higher GPAs on average. Therefore, when making decisions on potential students' admission status, the admissions office should be looking for scores in the medium to very high range. Students in the very low to low range may be at a higher risk for lower GPAs and potentially dropping out of classes due to difficulty. Universities and other education institutes that have very limited enrollment could use the Grit-V to help facilitate the decision-making process. For example, when making a decision between two students, both with equally strong academic track records but significantly different grit levels, the program may decide in favor of the student with higher grit since this student may be more likely to succeed.

Concerning admissions decisions, it is important to note that the results from this study do not make connects between low grit and failure. Having low grit doesn't necessarily mean high susceptibility to failure. For example, it is possible to have a high GPA, succeed in your

degree program, but have relatively low grit. Having a high GPA outweighs low grit in terms of academic success.

The study findings do not propose that the Grit-V be used in the same way as standardized tests in aiding in the admissions decision-making process. Rather, Grit-V could be added to the supplemental materials that many admissions offices use when making their decisions. The results of the Grit-V could perhaps be a tipping point when deciding on otherwise very similar candidates.

Limitations and Future Directions

Grit and Degree Completion

One limitation of this study was that data on degree completion was not collected. Degree completion would be an additional measure of academic success that would have a more long-term focus than the other variables used to measure academic success. With the addition of degree completion, researchers could get a better picture of how grit affects students' long-term goals, since graduation is the more likely desired result. If the Grit-V can be used to help graduate schools get a better picture of potential students' likelihood of finishing the program, that would be one more use of the Grit-V.

Grit and Other Student Populations

This study focused on undergraduate students for the target population when researching the development of Grit-V. It could be beneficial for future research to investigate if the Grit-V functions the same for graduate students as it does for undergraduate students. It would also be interesting to see if the dimensionality of grit is different for graduate students. The Grit-V could also be used for high school students. The Grit-S is the current measurement instrument being used in grit research in high school settings (e.g., Ivcevic & Brackett, 2014; Muenks, Wigfield,

Yang, & O'Neal, 2017). Since this study showed that the Grit-V provided more evidence of certain types of validity than the Grit-S, future researchers are encouraged to consider using Grit-V when selecting grit assessment measures.

Longitudinal Grit Research

Another limitation of this study was the use of cross-sectional data. By using this type of data, only the correlations between grit and academic success were studied at one point in time. Future researchers of grit are encouraged to explore the possibilities of using longitudinal measurement of grit. This would provide repeated observations of the same subjects over a period of time. We could then learn about growth patterns of grit. Does grit fluctuate throughout students' academic career? Do the interactions with grit and other non-cognitive skills change over time? Can grit be significantly increased through intervention? How does social environment affect one's grit? Should grit still be considered as a trait, as Duckworth claims? If not, how should it be categorized? These are the types of questions that could be answered with longitudinal data that are not possible with cross sectional data.

Grit and Self-regulation

This research did not examine correlations between grit and self-regulation. It would be advantageous to explore this type of relationship since little has been done on this topic. There has been research connecting self-regulation and success in school (e.g., Kitsantas, Winsler, & Huie, 2008; Duckworth & Carlson, 2013; Nicol, 2006), little has been done in making direct connections between grit and self-regulation in academic settings. Understanding this relationship will shed more light on grit and self-regulation as constructs, as well as how those constructs interact with each other.

Gender Differences

The slight difference in grit between males and females is statistically significant, albeit very slight. I would direct future researchers who are interested in grit and gender to explore this further as there does not appear to be a consensus on grit and gender interaction; Duckworth and Quinn (2009) concluded that there is no statistically significant difference between males' and females' grit level, while Rojas, Rese, Usher, and Toland (2012) found that girls reported higher levels of grit than boys. Perhaps a larger sample of participants would provide more meaningful results. It would also be crucial to explore why females are slightly grittier than males, if this is the case.

Grit and SES

This study provided evidence that SES was not a moderator variable between grit and GPA. However, study findings indicated that typically parents with higher education level had children with higher grit. Sirin's (2005) meta-analysis review suggested that parents' SES had a strong impact on students' academic achievement, and Huang and Zhu (2017) found that perceived level of grit among students can reduce the achievement gap between varying SES levels. Further research into this area could shine more light on how a people's upbringing affects their grit levels. This could be done either qualitatively (perhaps through ethnographic or focus group research) or quantitatively (such as investigating the relationships between grit and education level).

Grit and Children

The current trend in grit research focuses on adults. However, it may be worthwhile to investigate grit in children as well. Tough (2013) challenged the notion of cognitive skills being the main characteristic of successful children, but he had simply started the conversation.

Further research can help build upon his theory that non-cognitive skills such as grit and curiosity are very telling of a child's potential to be successful. From a measurement perspective, it could be valuable to develop a grit scale specifically designed as a diagnostic tool for children. If grit is an indicator of a child's potential for success, then developing an assessment to capture this could be vital in the field of childhood development. Since non-cognitive skills like grit are malleable and able to be developed, having a measurement tool that can assess grit in children can assist teachers in targeting students who need improvement in these areas.

Socially Desirable Responses

One of the caveats of any self-report scale is the potential for socially desired responses. These socially desired responses, or response biases, are "any systematic tendency to answer questionnaire items on some basis that interferes with accurate self-reports" (Braun, 2002, p. 49). It is a real possibility that the participants in this study selected responses that they deemed as more socially desirable than others. This can ultimately lead to participants providing less valid responses. In the item development process, research was carefully conducted to write items in a way to reduce this possibility; nonetheless, it may not be possible to fully eliminate response bias. The researcher also made it aware to the participants that they should provide frank answers, and that none of the responses would have any impact on their well-being. However, this would be much more difficult, if not impossible to control for, if the scale were to be used as supplemental material for admissions.

The Use of Grit Measurement in Business

Beyond college admissions, grit can be used as potential hiring criteria as well. Why would businesses be interested in grit? A gritty individual is more likely to pursue long-term goals and actually accomplish them. Having employees who possess the drive and willingness to set goals

and finish them is valuable in any environment. Businesses place a premium on non-cognitive skills such as agreeableness, conscientiousness, and grit (Brunello & Schlotter, 2011). These can be considered as soft skills (a culmination of personality traits, motivations, and goals that are valuable in multiple arenas, including business and school) and they have been shown to both predict and produce success in our lives (Heckman & Kautz, 2012). Achievement tests do not capture the essence of these non-cognitive skills. Thus, grit can offer a solution to this problem. Even though the vignettes in the Grit-V revolve around academic contexts, it is possible to develop a new series of vignettes that revolve around business situations.

The Use of Vignettes in Measurement

This study provided evidence that using vignettes in measurement can provide a more valid form of measurement than simple short-sentence item endorsement. Testing companies such as ETS and Pearson, as well as researchers interested in scale development, are encouraged to consider using vignettes when developing assessments. Vignettes add a layer of context that is simply not possible in simple sentence items (Alexander & Becker, 1978). The major downside of this technique is resources; the process of collecting focus group data, follow-up interviews with participants, and expert-panel review is a lengthy process. It also requires more time to take the assessment because participants must read longer texts (however, it may be possible to capture the same or more amount of information from fewer vignettes than from short sentence items). Future research efforts could start by developing vignette-based forms of assessment for other popular non-cognitive skills among researchers (e.g., self-efficacy, hope, optimism, and motivation).

Final Remarks

The purpose of this study was to investigate a new method of item development for the non-cognitive construct grit through the use of vignettes. This was not an attempt to discredit use of Grit-S, nor to say that all assessments that use simple short-sentence item endorsements are inherently bad forms of measurement. Rather, scale developers and researchers are strongly encouraged to consider vignettes when creating new assessments, as well as, adapting existing scales to be more current non-cognitive skills like grit, which are not easily observable. Therefore, we need these forms of assessment to measure the grittiness of an individual.

Measuring and defining academic success has been and will likely continue to be a heated debate among researchers, educators, and policymakers. What can be agreed upon though is that academic success is a vital part of education as a whole. While grit can be an excellent tool in getting a more rounded picture of an individual's potential for success in school, it is probably not wise to place too much emphasis on non-cognitive skills. The grit assessment is to be used as supplemental material, not as a replacement for existing measures or high stakes testing. It is the responsibility of everyone involved in education to make sure these types of assessment are used and interpreted appropriately. The findings and interpretations provided in this study are meant to be a starting point, and like non-cognitive skills themselves, are malleable and able to change with time.

REFERENCES

- Alexander, C. S., & Becker, H. J. (1978). The use of vignettes in survey research. *Public Opinion Quarterly*, 42(1), 93-104.
- Allen, D. (1999). Desire to finish college: An empirical link between motivation and persistence. *Research in Higher Education*, 40(4), 461-485.
- American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). (2014). *Standards for Educational and Psychological Testing*. Washington, DC: AERA.
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43, 561-73.
- Atkinson, J. W. (1964). *An introduction to motivation*. Oxford, England: Van Nostrand.
- Arthur, J., Kristjánsson, K., & Thoma, S. (2016). Is grit the magic elixir of good character? *Virtue Insight*.
- Austin, E. J., Deary, I. J., Gibson, G. J., McGregor, M. J., & Dent, J. B. (1998). Individual response spread in self-report scales: Personality correlations and consequences. *Personality and Individual Differences*, 24(3), 421-438.
- Babbie, R. R. 1995. *The practice of social research*. New York: Wadsworth.
- Bago d'Uva, T., Van Doorslaer, E., Lindeboom, M., & O'donnell, O. (2008). Does reporting heterogeneity bias the measurement of health disparities? *Health Economics*, 17(3), 351-375.
- Baer, J., Baldi, S., & Merola, S.S. (2005). *The measurement of socioeconomic status (SES) in the social sciences: A review to inform NAEP*. Unpublished paper prepared for the National Center of Education Statistics.
- Barbour, R. (2007). *Doing focus groups*. Thousand Oaks: Sage.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1-26.
- Basch, C. E. (1987). Focus group interview: an underutilized research technique for improving theory and practice in health education. *Health Education Quarterly*, 14(4), 411-448.
- Bebeau, M. J., & Thoma, S. J. (1999). "Intermediate" concepts and the connection to moral education. *Educational Psychology Review*, 11(4), 343-360.

- Beghtol, C. (1995). Mapping sentences and classification schedules as methods of displaying facets. *Advances in Classification Research Online*, 6(1), 1-12.
- Berque, D. A., Prey, J., & Reed, R. H. (Eds.). (2006). *The impact of tablet PCs and pen-based technology on education: Vignettes, evaluations, and future directions*. Purdue University Press.
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). New York, NY: Routledge.
- Borg, I., & Shye, S. (1995). *Facet theory: Form and content*. Sage.
- Braun, H. I., Jackson, D. N., & Wiley, D. E. (2001). Socially desirable responding: The evolution of a construct. In *The role of constructs in psychological and educational measurement* (pp. 61-84). Routledge.
- Brondani, M. A., MacEntee, M. I., Bryant, S. R., & O'Neill, B. (2008). Using written vignettes in focus groups among older adults to discuss oral health as a sensitive topic. *Qualitative Health Research*, 18(8), 1145-1153.
- Brown, D. G. (2000). *Interactive Learning: Vignettes from America's Most Wired Campuses*. Bolton, MA: Anker Publishing Company.
- Brunello, G., and Schlotter, M., (2011) *Non-cognitive skills and personality traits: Labor market relevance and their development in education and training systems* (IZA Discussion Paper No. 5743). Retrieved from <http://ftp.iza.org/dp5743.pdf>
- Button, S. B., Mathieu, J. E., & Zajac, D. M. (1996). Goal orientation in organizational research: A conceptual and empirical foundation. *Organizational Behavior and Human Decision Processes*, 67, 26-48.
- Calder, B. J. (1977). Focus groups and the nature of qualitative marketing research. *Journal of Marketing research*, 353-364.
- Campbell, P. B. (1996). How would I handle that? Using vignettes to promote good math and science education. Washington, DC: American Association for the Advancement of Science.
- Carter, N. L., Guan, L., Maples, J. L., Williamson, R. L., & Miller, J. D. (2015). The downsides of extreme conscientiousness for psychological wellbeing: The role of obsessive compulsive tendencies. *Journal of Personality*. doi:10.1111/jopy.12177.
- Carver, C. S., Scheier, M. F., & Segerstrom, S. C. (2010). Optimism. *Clinical Psychology Review*, 30(7), 879-889.
- Chapin, F.S. (1933). *The measurement of social status*, 3, University of Minnesota Press.

- Chau, J. P. C., Chang, A. M., Lee, I. F. K., Ip, W. Y., Lee, D. T. F., & Wootton, Y. (2001). Effects of using videotaped vignettes on enhancing students' critical thinking ability in a baccalaureate nursing programme. *Journal of Advanced Nursing*, *36*(1), 112-119.
- Cofer, C. N., & Appley, M. H. (1964). *Motivation: Theory and research*. Oxford, England: John Wiley.
- Colquitt, J. A., & Simmering, M. J. (1998). Conscientiousness, goal orientation, and motivation to learn during the learning process: A longitudinal study. *Journal of Applied Psychology*, *83*(4), 654.
- Costa, P. T., Jr., & McCrae, R. R. (1992). *The NEOPI-R professional manual*. Odessa, FL: Psychological Assessment Resources.
- Cross, T. M. (2014). The gritty: Grit and non-traditional doctoral student success. *Journal of Educators Online*, *11*(3), n3.
- Datu, J., Yuen, M., & Chen, G. (2017). Grit and determination: A Review of literature with implications for theory and research. *Journal of Psychologists and Counsellors in Schools*, *27*(2), 168-176. doi:10.1017/jgc.2016.2
- Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, *5*(4), 194-197.
- Deci, E., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Springer Science & Business Media.
- Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist*, *26*(3-4), 325-346.
- Deci, E. L., & Ryan, R. M. (Eds.). (2002). *Handbook of self-determination research*. Rochester, NY: University Rochester Press.
- Deci, E. L., & Ryan, R. M. (2010). *Self-determination*. Ney York, NY: John Wiley & Sons, Inc.
- DeSimone, J. A., Harms, P. D., & DeSimone, A. J. (2015). Best practice recommendations for data screening. *Journal of Organizational Behavior*, *36*(2), 171-181.
- DeVellis, R. F. (2016). *Scale development: Theory and applications* (Vol. 26). Thousand Oaks, CA: Sage.
- Diener, C. I., & Dweck, C. S. (1978). An analysis of learned helplessness: Continuous changes in performance, strategy, and achievement cognitions following failure. *Journal of Personality and Social Psychology*, *36*, 451– 462.
- Diener, C. I., & Dweck, C. S. (1980). An analysis of learned helplessness: II. The processing of success. *Journal of Personality and Social Psychology*, *39*, 940 –952.

- Dowd, J. B., & Todd, M. (2011). Does self-reported health bias the measurement of health inequalities in US adults? Evidence using anchoring vignettes from the Health and Retirement Study. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 66(4), 478-489.
- Downing, S. M., & Haladyna, T. M. (1997). Test item development: Validity evidence from quality assurance procedures. *Applied Measurement in Education*, 10(1), 61-82.
- Dresselhaus, T. R., Peabody, J. W., Luck, J., & Bertenthal, D. (2004). An evaluation of vignettes for predicting variation in the quality of preventive care. *Journal of General Internal Medicine*, 19(10), 1013-1018.
- Duckworth, A. (2016). *Grit: The power of passion and perseverance*. New York, NY: Simon and Schuster.
- Duckworth, A.L., & Yeager, D.S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher*, 44(4), 237-251.
- Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (GRIT-S). *Journal of Personality Assessment*, 91(2), 166-174.
- 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.
- Duckworth, A. L., & Seligman, M. E. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science*, 16(12), 939-944.
- Duckworth, A. L., & Seligman, M. E. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology*, 98(1), 198.
- Duckworth, A. L., & Carlson, S. M. (2013). Self-regulation and school success. *Self-regulation and Autonomy: Social and Developmental Dimensions of Human Conduct*, 40, 208.
- Duckworth, A. L. (2013) *Grit: The power of passion and perseverance* [Video file]. Retrieved from https://www.ted.com/talks/angela_lee_duckworth_grit_the_power_of_passion_and_perseverance
- 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.
- Duckworth, A. L., Weir, D. R., Tsukayama, E., & Kwok, D. (2012). Who does well in life? Conscientious adults excel in both objective and subjective success. *Frontiers in Psychology*, 3, 356.
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, 41, 1040-1048.

- Dweck, C. S. (1989). Motivation. In A. Lesgold & R. Glaser (Eds.), *Foundations for a Psychology of Education*, (pp. 87-136). Hillsdale, NJ: Erlbaum.
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press.
- Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York, NY: Random House Incorporated.
- Dweck, C. (2015). Carol Dweck revisits the 'growth mindset'. *Education Week*, 35(5), 20-24.
- Dweck, C. S., Hong, Y., & Chiu, C. (1993). Implicit theories: Individual differences in the likelihood and meaning of dispositional inference. *Personality and Social Psychology Bulletin*, 19, 644-656.
- Elliot, R., Fischer, C.T., & Rennie, D.L. (1999). Evolving guidelines for publication of qualitative research studies in psychology and related fields. *British Journal of Clinical Psychology*, 38(3), 215-229.
- Engelhard Jr, G (2013). *Invariant measurement: Using Rasch models in the social, behavioral, and health sciences*. Chicago, IL: Routledge.
- Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100, 363–406.
<http://dx.doi.org/10.1037/0033-295X.100.3.363>_Eskreis-
- Fan, X., Miller, B. C., Park, K. E., Winward, B. W., Christensen, M., Grotevant, H. D., & Tai, R. H. (2006). An exploratory study about inaccuracy and invalidity in adolescent self-report surveys. *Field Methods*, 18(3), 223-244.
- Farmer, R., & Sundberg, N. D. (1986). Boredom proneness--the development and correlates of a new scale. *Journal of Personality Assessment*, 50(1), 4-17.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Thousand Oaks, CA: Sage.
- Fujii, E. T., Hennessy, M., & Mak, J. (1985). An evaluation of the validity and reliability of survey response data on household electricity conservation. *Evaluation Review* 9 (1): 93–104.
- Hanushek, E. A., & Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth*, 17(4), 267-321.
- Haladyna, T. M., & Downing, S. M. (Eds.). (2011). *Handbook of test development*. New York, NY: Routledge.
- Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items*. New York, NY: Routledge.

- Hanson, S. L. (1994). Lost talent: Unrealized educational aspirations and expectations among U.S. youths. *Sociology of Education*, 64, 263–277.
- Hartigan, J., & Wigdor, A. (1989). Fairness in employment testing. *Science*, 245(4913), 14-14.
- Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, 19(4), 451-464.
- Herold, B. (2015). Is 'grit' racist. *Digital education*. Retrieved from http://blogs.edweek.org/edweek/DigitalEducation/2015/01/is_grit_racist.html.
- Hochanadel, A., & Finamore, D. (2015). Fixed and growth mindset in education and how grit helps students persist in the face of adversity. *Journal of International Education Research*, 11(1), 47.
- Huang, H., & Zhu, H. (2017). High achievers from low socioeconomic backgrounds: The critical role of disciplinary climate and grit. *Mid-Western Educational Researcher*, 29(2).
- Hughes, R. (2012). Vignettes. In L. M. Given (Ed.), *The SAGE Encyclopedia of Qualitative Research Methods* (pp. 919-920). DOI: <http://dx.doi.org/10.4135/9781412963909>
- Hughes, R., & Huby, M. (2002). The application of vignettes in social and nursing research. *Journal of Advanced Nursing*, 37(4), 382-386.
- Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior*, 26(4), 331-362.
- Goldberg, L. R. (1990). An alternative" description of personality": the big-five factor structure. *Journal of Personality and Social Psychology*, 59(6), 1216.
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48(1), 26.
- Gorman, R. (2015). *An examination of academic grit in urban high schools* (Doctoral dissertation). Retrieved from <http://scholarworks.wmich.edu/dissertations/1174>.
- Graziano, W. G., & Eisenberg, N. (1997). Agreeableness: A dimension of personality. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 795-824). San Diego, CA: Academic Press.
- Griffin, B., & Hesketh, B. (2004). Why openness to experience is not a good predictor of job performance. *International Journal of selection and assessment*, 12(3), 243-251.
- Grol-Prokopczyk, H., Freese, J., & Hauser, R. M. (2011). Using anchoring vignettes to assess group differences in general self-rated health. *Journal of Health and Social Behavior*, 52(2), 246-261.

- Gumpel, T., & Wilson, M. (1996). Application of a Rasch analysis to the examination of the perception of facial affect among persons with mental retardation. *Research in Developmental Disabilities, 17*(2), 161-171.
- Ivcevic, Z., & Brackett, M. (2014). Predicting school success: Comparing conscientiousness, grit, and emotion regulation ability. *Journal of Research in Personality, 52*, 29-36.
- Jernigan, C. G. (2004). What do students expect to learn? The role of learner expectancies, beliefs, and attributions for success and failure in student motivation. *Current Issues in Education, 7*(4).
- Jeronimus, B. F., Ormel, J., Aleman, A., Penninx, B. W., & Riese, H. (2013). Negative and positive life events are associated with small but lasting change in neuroticism. *Psychological Medicine, 43*(11), 2403–15.
doi:10.1017/s0033291713000159.
- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and Research, 2*(1999), 102-138.
- Kamberelis, G., & Dimitriadis, G. (2005). Focus groups: Strategic articulations of pedagogy, politics, and inquiry. In N. K. Denzin & Y. Lincoln (Eds.), *The Sage Handbook of Qualitative Research, 3rd ed.* (pp. 887-907). Thousand Oaks, CA: Sage.
- Kautz, T., Heckman, J. J., Diris, R., Ter Weel, B., & Borghans, L. (2014). *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success* (No. w20749). Cambridge, MA: National Bureau of Economic Research.
- King, G., Murray, C.J. L., Salomon, J.A., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review, 98*(1), 191-207.
- Kitsantas, A., Winsler, A., & Huie, F. (2008). Self-regulation and ability predictors of academic success during college: A predictive validity study. *Journal of Advanced Academics, 20*(1), 42-68.
- Komaraju, M., Ramsey, A., & Rinella, V. (2013). Cognitive and non-cognitive predictors of college readiness and performance: Role of academic discipline. *Learning and Individual Differences, 24*, 103-109.
- Komaraju, M., Karau, S. J., & Schmeck, R. R. (2009). Role of the Big Five personality traits in predicting college students' academic motivation and achievement. *Learning and individual differences, 19*(1), 47-52.
- Kyllonen, P.C., & Bertling, J. (Eds.). (2013) Innovative questionnaire assessment methods to increase cross-country comparability. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *A Handbook of International Large-Scale Assessment Data Analysis:*

- Background, Technical Issues, and Methods of Data Analysis* (pp. 277-287). London, UK: Chapman Hall/CRC Press.
- Linacre, J. M. (1989). Multi-faceted Rasch measurement. Chicago, IL: MESA Press.
- Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, 16 (2), 878.
- Linacre, J. M. (2015). Facets Rasch Measurement (Version 3.71.4). Chicago, IL: Winsteps.com.
- Lincoln, Y.S., & Guba, E.G. (1985). *Naturalistic Inquiry*. Newbury Park, CA: Sage.
- MacCann, C., Duckworth, A. L., & Roberts, R. D. (2009). Empirical identification of the major facets of conscientiousness. *Learning and Individual Differences*, 19(4), 451-458.
- MacIntosh, J. A. (1993). Focus groups in distance nursing education. *Journal of Advanced Nursing*, 18(12), 1981-1985.
- Marks, H.M. (2000). Student engagement in instructional activity: patterns in the elementary, middle, and high school years. *American Educational Research Journal*, 37(1), 152-184.
- Martin, E. A., Campanelli, P. C., & Fay, R. E. (1991). An application of Rasch analysis to questionnaire design: Using vignettes to study the meaning of work in the current population survey. *The Statistician*, 265-276.
- McCrae, R. R., & Costa, P. T, Jr. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52, 81-90.
- Miller, L.A., & Lovler, R.L. (2016). *Foundations of psychological testing: A practical approach*. Thousand Oaks, CA: Sage.
- Muenks, K., Wigfield, A., Yang, J. S., & O'Neal, C. R. (2017). How true is grit? Assessing its relations to high school and college students' personality characteristics, self-regulation, engagement, and achievement. *Journal of Educational Psychology*, 109(5), 599.
- Nicol, D. (2006, December). *Increasing success in first year courses: Assessment re-design, self-regulation and learning technologies*. In proceedings of the 23rd annual ASCILITE conference, Sydney, Australia.
- Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *School Field*, 7(2), 133-144.
- Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology*, 93(1), 116.

- Norcini, J. (2004). Back to the future: Clinical vignettes and the measurement of physician performance. *Annals of Internal Medicine, 141*(10), 813-814.
- Nosanchuk, T. A. (1972). The vignette as an experimental approach to the study of social status: An exploratory study. *Social Science Research, 1*(1), 107-120.
- O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual Differences, 43*(5), 971-990.
- Ormrod, J. E. (2003). *Educational psychology, developing learners*. Upper Saddle River, NJ: Merrill Prentice Hall.
- Peabody, J. W., Luck, J., Glassman, P., Jain, S., Hansen, J., Spell, M., & Lee, M. (2004). Measuring the quality of physician practice by using clinical vignettes: A prospective validation study. *Annals of Internal Medicine, 141*(10), 771-780.
- Peabody, J. W., Luck, J., Glassman, P., Dresselhaus, T. R., & Lee, M. (2000). Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *Jama, 283*(13), 1715-1722.
- Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: the use of factor analysis for instrument development in health care research*. Thousand Oaks, CA: Sage.
- Perkins-Gough, D. (2013). The significance of grit: A conversation with Angela Lee Duckworth. *Educational Leadership, 71*(1), 14-20.
- Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1*(3), 385-401.
- Randall, J., & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education, 26*(7), 1372-1380.
- Randall, J., & Engelhard, J. G. (2010). Using Guttman's mapping sentences and Many Facet Rasch Measurement Theory to develop an instrument that examines the grading philosophies of teachers. *Journal of Applied Measurement, 11*(2), 122-141.
- Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Achievement Tests* (Expanded edition, 1980). Chicago, IL: University of Chicago Press.
- Reckase, M. D. (1979). Unifactor Latent trait models applied to multifactor tests: Results and implications. *Journal of Educational and Behavioral Statistics, 4*(3), 207-230. <https://doi.org/10.3102/10769986004003207>
- Reckase, M. (2009). *Multidimensional item response theory* (Vol. 150). New York, NY: Springer.

- Roach, L. E., & Wandersee, J. H. (1993). Short story science: Using historical vignettes as a teaching tool. *Science Teacher*, 60(6), 18-21.
- Rojas, J. P., Reser, J. A., Usher, E. L., & Toland, M. D. (2012). Psychometric properties of the academic grit scale. *Lexington: University of Kentucky*.
- Rosenthal, R. & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data-analysis*. New York: McGraw-Hill.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68.
- Salomon, J. A., Tandon, A., & Murray, C. J. (2004). Comparability of self rated health: cross sectional multi-country survey using anchoring vignettes. *BMJ*, 328(7434), 258.
- Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: assessment and implications of generalized outcome expectancies. *Health Psychology*, 4(3), 219.
- Schumacker, R. (2015). *Using R with multivariate statistics*. Thousand Oaks, CA: Sage.
- Seibert, S. E., & Kraimer, M. L. (2001). The five-factor model of personality and career success. *Journal of Vocational Behavior*, 58(1), 1-21.
- Sherer, M., Maddux, J. E., Mercandante, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R. W. (1982). The self-efficacy scale: Construction and validation. *Psychological Reports*, 51(2), 663-671.
- Sims, V.M. (1927). *The measurement of socioeconomic status*. Bloomington, IL: Public School Printing Co.
- Sirin, S.R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research*, 75, 417–453.
- Shye, S., Elizur, D., & Hoffman, M. (1994). *Introduction to facet theory: Content design and intrinsic data analysis in behavioral research*. Thousand Oaks, CA: Sage.
- Snyder, C. R., Harris, C., Anderson, J. R., Holleran, S. A., Irving, L. M., Sigmon, S. T., Yoshinobu, L, Gibb, J., Langelle, C. &Harney, P. (1991). The will and the ways: Development and validation of an individual differences measure of hope. *Journal of personality and social psychology*, 60(4), 570.
- Snyder, C. R. (1994). *The psychology of hope: You can get there from here*. New York, NY: Simon and Schuster.
- Snyder, C. R., Shorey, H. S., Cheavens, J., Pulvers, K. M., Adams III, V. H., & Wiklund, C. (2002). Hope and academic success in college. *Journal of Educational Psychology*, 94(4), 820.

- Smith, R. M. (2004). Fit analysis in latent trait models. In E. V. Smith & R. M. Smith (Eds.), *Introduction to Rasch Measurement* (pp. 73–92). Maple Grove, MN: JAM Press.
- Spalding, N. J. (2004). Using vignettes to assist reflection within an action research study on a preoperative education programme. *British Journal of Occupational Therapy*, 67(9), 388-395.
- 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.
- Stewart, D. W., & Shamdasani, P. N. (2014). *Focus groups: Theory and practice* (Vol. 20). Sage.
- Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York, NY: Cambridge University Press.
- Stoll-Kleemann, S., O’Riordan, T., & Jaeger, C. C. (2001). The psychology of denial concerning climate mitigation measures: evidence from Swiss focus groups. *Global Environmental Change*, 11(2), 107-117.
- Stosny, S. (2011). Self-regulation. *Psychology Today*. Retrieved from <https://www.psychologytoday.com/us/blog/anger-in-the-age-entitlement/201110/self-regulation>
- Sulea, C., Van Beek, I., Sarbescu, P., Virga, D., & Schaufeli, W. B. (2015). Engagement, boredom, and burnout among students: Basic need satisfaction matters more than personality traits. *Learning and Individual Differences*, 42, 132-138.
- Taber, K. S. (2018). The use of Cronbach’s alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48(6), 1273-1296.
- Taussig, F.W. (1920). *Principles of economics*. Newcastle: Cambridge Scholars Publishing.
- Teahan, J. E. (1958). Future time perspective, optimism, and academic achievement. *The Journal of Abnormal and Social Psychology*, 57(3), 379.
- Terman, L. M., Oden, M. H., & Bayley, N. (1947). *The gifted child grows up: Twenty-five years'follow-up of a superior group*. Stanford, CA: Stanford University Press.
- Thompson, E.R. (2008). Development and Validation of an International English Big-Five Mini-Markers. *Personality and Individual Differences*, 45(6), 542–548.
- Tilden, V. P., Nelson, C. A., & May, B. A. (1990). Use of qualitative methods to enhance content validity. *Nursing Research*, 39, 172-175.
- Tough, P. (2013). *How children succeed: Grit, curiosity, and the hidden power of character*. London, England: RH Books.

- Tracey, T. J., & Sedlacek, W. E. (1987). Prediction of college graduation using noncognitive variables by race. *Measurement and Evaluation in Counseling and Development*, 19(4), 177-184.
- Tracey, T. J., & Sedlacek, W. E. (1989). Factor structure of the non-cognitive questionnaire-revised across samples of Black and White college students. *Educational and Psychological Measurement*, 49(3), 637-648.
- Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., & Vallieres, E. F. (1992). The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement*, 52(4), 1003-1017.
- Von Culin, K. R., Tsukayama, E., & Duckworth, A. L. (2014). Unpacking grit: Motivational correlates of perseverance and passion for long-term goals. *The Journal of Positive Psychology*, 9(4), 306-312.
- Wilkinson, S. (1998). Focus groups in health research: Exploring the meanings of health and illness. *Journal of Health Psychology*, 3(3), 329-348.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. New York, NY: Taylor & Francis.
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, 34(6), 806-838.
- Wright, B. D. (1996). Reasonable mean-square fit values. In Wright, B. D., & Linacre, J. M., (Eds.) *Rasch Measurement Transactions. Part 2*. Chicago: MESA Press.
- Wu, M., & Adams, R. J. (2013). Properties of Rasch residual fit statistics. *Journal of Applied Measurement*, 14(4), 339-355.
- Yazzie-Mintz, E. (2007). *Voices of students on engagement: A report on the 2006 high school survey of student engagement*. Indiana: Center for Evaluation and Education Policy.
- Zepke, N., & Leach, L. (2010). Beyond hard outcomes: 'Soft' outcomes and engagement as student success. *Teaching in Higher Education*, 15(6), 661-673.

APPENDIX A:
FOCUS GROUP GUIDE

Hello _____, my name is _____, it's nice to [meet you/see you all again]. Thank you for taking the time to talk with me this [afternoon/morning]. Before we get started, I want to take a minute to review the purpose of this focus group: I am interested in hearing about your perspective on grit. Grit is a combination of perseverance and consistency of interest, and it's a personality trait that can help a person achieve their long-term goals. I am interested in hearing about your experiences when you were, or were not, gritty. The information we gather will be used to help improve my understanding of grit, as well as help me write items for a vignette-based scale that aims to measure a person's grit. The data gathered from these focus groups will be used in my dissertation. Anything that you say will be kept strictly confidential. The audio file from this focus group will be transcribed to text and then destroyed within one year of the focus group time. The resulting text file will be de-identified. Please feel free to respond candidly and honestly. The focus group should not take any longer than 30 to 45 minutes.

I also want you to know that your participation in this focus group is entirely optional. You don't have to participate and there will be no penalty for not participating. If you decide you would like to start and then part of the way through change your mind, it is okay to stop at any time. During the focus group you may see me taking notes – these notes help keep me on track and ensure I don't repeat questions that I would like to ask. We are also recording this conversation. The recording will be deleted following transcription. In addition, if you say

something during the focus group and decide later that you do not want us to use it, we can redact the tape.

Does everything sound alright? [wait for response] Is it okay to begin? [wait for response] Do you have any questions about the focus group or any of the other information I have given to you before we begin? [wait for response] Okay, then let's begin.

Some advice for the focus group, try and wait for the person talking to finish then say your comments as it gets very confusing if everyone talks over the speaker. Perhaps raise your hand and we will come to you next.

[Note that questions in italics are planned questions to be asked during the focus group. All other questions are possible probing questions that may be employed based on the interviewee's response to initial questions]

Background Information – Let's begin by talking about your background. We will then move on to questions specifically about grit.

- *What is your understanding of grit?*
 - When you think of grit, what comes to mind?
 - Is there more to grit than perseverance and resilience?
- *Please tell me a little about your short term and long terms goals as a university student*
 - Were there any obstacles in your way?
 - How did you overcome them?
 - What made it too difficult to overcome?
- *Please describe a few situations when you persevered through a challenging obstacle to accomplish your goals*
 - What made those hardships particularly memorable?
- *Please describe a situation when you gave up on a goal because your interests changed?*
 - Did you find it difficult to continue pursuing your goal when something more exciting comes along?
- *What does a gritty person do that makes them gritty?*
 - What separates a gritty person and a non-gritty person?
 - Do you believe there are varying levels of grit, or are you either gritty or not?

- *Can you describe someone you know who believe to be a very gritty person? Parents? Family member? Friend? Coworker?*
 - What do they do that makes them gritty?
- *How important do you think it is to be gritty in terms of achieving your goals?*
 - Can you be too gritty?
 - Is being gritty enough?
- *Is there anything else you would like to tell me about?*

APPENDIX B:

GRIT-V SCALE (INITIAL ITEM POOL FOR PILOT STUDY)

Item 1: Consistency of Interest

Suzy has been interested in becoming a nurse for several years. She is in her third year of study and is beginning to grow frustrated with the coursework because it seems to be less fascinating than when she first started. Her roommate, a marketing major, shows Suzy some slides from her most recent lecture, and Suzy's interest is piqued. After some time, she decides to put nursing on hold and take a few marketing classes to see if she wants to change her field.

If you were in her position, what would you most likely do?

- 1) Since I'm in my third year, I would not let a random interest get in the way of my long-term passion to become a nurse.
- 2) Since I'm a little burned out on nursing, I may take a semester off, but ultimately I would return to finish my goal.
- 3) I would put nursing on hold and take a few marketing classes to see if I want to change fields.
- 4) Life is too short to pursue something you lose interest in. I would go to my advisor and ask how I can submit the paperwork to sign up for marketing as my major.

Item 2: Perseverance of effort

Robert is starting his third year as an engineering student at a university. So far, his studies have not been too demanding. However, this semester he is encountering several difficult courses that are required for his degree. They are so difficult in fact that he is questioning his ability to complete them with high enough grades to continue with his program.

If you were in his position, what would you most likely do?

- 1) I would not let the difficulties discourage me from pursuing my dream to be an engineer. I would fight through it and use all available resources, even if it means I will sacrifice something else important to me.
- 2) I would be slightly discouraged, but I would stick with it as long as I could. If I get to the midway point in the semester and I'm still struggling, I might drop one or more of the classes.

- 3) I would most likely drop one or two of the most difficult classes. If I still struggle, I would think about changing majors to something more attainable.
- 4) Life is too short to be constantly struggling. I would probably give up on engineering and pursue a different career choice.

Item 3: Consistency of Interest

Howard set himself a goal to complete his schoolwork before participating in the extracurricular activities offered at the college. However, the timing of the activities is right after class, and many of his friends are trying to persuade him that as long as he does his work, it's not a big deal. Howard knows himself too well, and if given the opportunity to procrastinate, he will.

If you were in his position, what would you most likely do?

- 1) I would just do my work first. It's not worth risking putting things off and possibly not doing it. If I set a goal, I finish it.
- 2) I would attend the extracurricular activities my friends keep talking about first to see if it's worth postponing my goal. I'll likely not change my mind though.
- 3) My friends may have a point. I would try postponing my goal to see if it has any impact on my studies before completely abandoning my goal.
- 4) My friends are right. I should be social and I can always get me work done later. It can't be that big of a deal to put off this small goal.

Item 4: Perseverance of Effort

Scott is not the best student, but he is hard working and can usually overcome difficulties. He is currently working with a group on a project, feels out of his element because the class is outside of his major, but the other group members are very familiar with the content. At times he feels out of place, but he doesn't want to change groups or work by himself.

If you were in Scott's position, how would you handle this?

- 1) I am a hard worker, so I would put in the extra time to become more familiar with the content so I can work more effectively with the group.
- 2) I typically work very hard, but since it's outside of my comfort zone it's a little harder for me to put in the extra time. I would still try though.
- 3) As much as I want to be useful in the group, I would probably end up being a burden as I find it hard to put in the needed effort when I'm not doing what I'm good at.
- 4) I would either switch groups or drop the class. As much as I like the class, I can't work hard at something that causes me grief.

Item 5: Consistency of Interest

Mary decided to take on multiple long-term projects at the same time. After making several commitments, she is feeling overwhelmed. Recently, a friend asked her to join yet another project, but this project seems much more appealing to her than her current one. She's going to have to choose one to focus on.

If you were in her position, what would you most likely do?

- 1) I made prior commitments, so I can't quit after this many months. I will finish these first, then take on new projects if I have time.
- 2) Even though I made commitments, I need to at least consider other projects even if that means I may end up abandoning one or more of them.
- 3) I would ask more about this new project because I'm strongly considering leaving some of my prior commitments because they no longer seem worthwhile.
- 4) I would take on the new project since it's new and refreshing.

Item 6: Consistency of interest

Mary has recently taken an interest in intermural sports. She enjoys it so much, that she decides to make becoming the team captain a long-term goal for her. She begins working toward becoming the team captain, but realizes later that it will require a huge time commitment. After a few months, she has almost reached her goal, but realized that her interest is tapering off. She doesn't know if she should continue or quit.

If you were in her position, what would you most likely do?

- 1) I finish what I begin. Even if I find it less appealing than before, I've spent too much time to quit now. I would finish what I started without hesitation.
- 2) I would chat with my family and friends to see what they recommend. Even though I find it less appealing than before, I would most likely continue until I was team captain.
- 3) Even though I would like to finish what I started, I'm not sure it's worth it. I would stop and pursue a new interest.
- 4) Since I find it less interesting, I would quit. I would rather pursue other interests since this one is no longer attractive to me.

Item 7: Consistency of Interest

Henry is in his third year of college. Every year, he has changed majors because he just can't seem to maintain his level of interest in his coursework. Now he is wondering if college is even the right path for him.

How much do you relate to Henry?

- 1) I relate to him very much. I find it very difficult to maintain interests over long periods of time.
- 2) I relate moderately. I struggle to maintain interests over long periods of time.
- 3) I don't relate very much. Sometimes I lose interest in things over time, but that doesn't often turn into me quitting or changing interests completely.
- 4) Not at all. I very rarely change my interests over long periods of time.

Item 8: Perseverance of Effort

Joe is having a very difficult time with his statistics course. His professor has given him many ways to get help, such as the assistance lab, free tutoring, and extended office hours. His professor also informed him that the class isn't required, so Joe doesn't need this class for his degree, but it could be useful later in his studies.

If you were in Joe's situation, what would you most likely do?

- 1) It's going to be difficult, but with diligence I am definitely capable of succeeding. I would stick with it.
- 2) Even though it's not required, I would stick with it until mid-semester, and if the extra help isn't helping I may quit.
- 3) Since it's not required, I am more likely to give up. It's hard for me to be diligent when it's not something necessary.
- 4) As soon as I found out it's something I don't need to do, I would stop. There is no point in being diligent when I don't have to be.

Item 9: Consistency of Interest

Diana is working on a project that she is very passionate about. Her friends even commented that she has become obsessed with it. However, after working on it for only a short time, she becomes bored of it and ultimately abandons it. This seems to happen to Diana fairly regularly that she loses interest in something she was interested in quickly.

How much do you relate to Diana?

- 1) I don't relate to her at all. When I become passionate, even obsessed with a project, I see it to the end. I don't lose interest over a short period of time.
- 2) I relate to her very little. Typically once I'm interested in something, I don't change my mind. It does happen once in a while though.

- 3) I relate to her. I find that my interests change from time to time over short periods of time.
- 4) I'm exactly like Diana. I change interests all the time, and always after a very short amount of time.

Item 10: Perseverance of Effort

Steven is in his last year of study, but is encountering a number of setbacks. It's getting to the point where he is considering delaying graduation, but if he does that it will create more problems for him. Most of the setbacks are related to his coursework.

If you were Steven, what would you most likely do?

- 1) I would suck it up and deal with the setbacks in every way imaginable so I can achieve my goal of graduating this year. I don't let setbacks weigh me down.
- 2) I would try to deal with the setbacks, but I understand I may not succeed and it is better for me to postpone graduating until next semester.
- 3) I would try to achieve my goal, but I probably wouldn't be able to deal with the setbacks and I would achieve my goal at a later time.
- 4) I do not deal with setbacks very well and would most likely not achieve my goal until a much later time.

Item 11: Consistency of Interest

John has been interested in finance since his freshman year. Now that he has finished his finance introduction classes, he doesn't seem to be as interested in the field as he once was. He takes a few engineering courses, and thinks that that may now be his calling. However, once he starts taking the tougher classes, he doesn't think engineering is his passion either. He simply can't stick to one academic path.

If you were John, what would you most likely do?

- 1) I tend to change interests frequently, so I would most likely drop engineering and start focusing on something else.
- 2) I may pursue one more engineering class to see if I really don't find it interesting anymore, but the odds are I would change majors.
- 3) I would stick with engineering a bit longer. I want to really make sure that I am not interested in it before making a change. I don't make changes lightly.
- 4) I would not change my major. I maintain my interests over long periods of time so I would not continue changing my majors.

Item 12: Perseverance of Effort

Joanne is working with a group of students on a class project. Normally she works very well in groups. However, this time she doesn't get along a few members, and is becoming a distraction. There is a lot on the line in regard to this project though, as the group with the best results will have their work showcased at an upcoming symposium for outstanding student work. This would be something great to include on her resume.

If you were Joanne, what would you most likely do?

- 1) If I can't work well with the group members, I can't put forth the effort needed to succeed. I would do enough to get by and try again next time.
- 2) As much as I want to win, it's too much effort if it's a struggle to work with my other members. I would put in some effort, but I wouldn't plan on winning.
- 3) While I know I may not win, I would try my best because this is an important thing to work toward. I would put in enough effort in order to give us a fighting chance to win.
- 4) I would put in my best effort best I need to win this contest. Nothing gets in my way when I have a goal, and working with difficult group members is no different.

Item 13: Consistency of Interest

Annabelle has always had problems with staying interested in her goals for long periods of time. Whenever she finds something that interests her, something else comes along and distracts her. Lately this has begun affecting her work because she starts projects without finishing them because she loses interests halfway through completing them.

How much do you relate to Annabelle?

- 1) I relate very much. I can't seem to stay interested in my goals for very long.
- 2) I somewhat relate to Annabelle. If I'm really passionate about something, I'm more likely to finish, but it is quite easy for me to get distracted.
- 3) I don't relate to Annabelle. If I start something, I try to finish. There may be times where I get distracted, but it's rare.
- 4) I don't relate at all. I finish everything I start, even if I lose interest over time. I must finish things I start before taking on something new.

Item 14: Perseverance of Effort

Karl has decided to add a minor to his degree plan. Several of his professors have advised him that a foreign language minor could help set him apart in the job market, so he has started studying Spanish. He is unfortunately finding it much more difficult than he had expected. It

appears that he will need to put a significant amount of time into the coursework, something he had not planned on.

If you were Karl, what would you most likely do?

- 1) I am a hard worker, and I would spend as much time as needed to get the most out of the class, even if I have to sacrifice my time elsewhere. I want to get the most of out it.
- 2) I would put in the effort, but I'm not sure I would sacrifice time elsewhere. It's important to me, but as long as I get the grades I need I'm happy.
- 3) I might put in enough time to finish the minor, but if it's a lot more work than I anticipated it may not be worth it and I could drop it.
- 4) Since it's so much more of a time commitment that I first thought, I would probably change my mind on getting a minor and focus on my other courses.

Item 15: Perseverance of Effort

Larry always changes his interests, and when that happens, he doesn't put forth his best effort. One day, he finds computer science fascinating, but the next day he loses interest. This carries over to his choices in his program of study, and as a result his counselor has suggested taking a "general education" tract until he is able to make a firmer decision on his long term goals.

How much do you relate to Larry?

- 1) Not at all. Even if I have other interests, they don't deter me and cause me to work less hard. I don't picture myself not being able to decide on a long-term goal and sticking with it.
- 2) I don't really relate to him. At times I struggle to stay interested in my subject, but ultimately I want to accomplish my goals, and working hard is the only way to accomplish them.
- 3) I somewhat relate to Larry. I too can't seem to decide on an academic track to follow. Even though deep down I know I shouldn't waver, I end up changing my goals frequently.
- 4) I relate to him a lot. No matter what I find interesting, it's only a matter of time before something else grabs my attention. I can't seem to find the right subject to study in school.

APPENDIX C:

GRIT-V SCALE (FINAL FORM)

Item 1: Consistency of Interest

Suzy has been interested in becoming a nurse for several years. She is in her third year of study and is beginning to get frustrated with the coursework because it seems to be less fascinating than when she first started. Her roommate, a marketing major, shows Suzy some slides from her most recent lecture, and Suzy's interest is piqued. After some time, she decides to put nursing on hold and take a few marketing classes to see if she wants to change her field.

If you were in her position, what would you most likely do?

- 1) Life is too short to pursue something you lose interest in. I would go to my advisor and ask how I can submit the paperwork to sign up for marketing as my major.
- 2) Since I'm starting to slowly lose interest in nursing, I may take a semester off, but ultimately I would return to finish my goal.
- 3) Since I'm in my third year, I would not let a random interest get in the way of my long-term passion to become a nurse.
- 4) I think keeping a balanced checkbook is more important than staying active and fit.

Item 2: Perseverance of effort

Robert is starting his third year as an engineering student at a university. So far, his studies have not been too demanding. However, this semester he is encountering several difficult courses that are required for his degree. They are so difficult in fact that he is questioning his ability to complete them with high enough grades to continue with his program.

If you were in his position, what would you most likely do?

- 1) I would not let the difficulties discourage me from pursuing my dream to be an engineer. I would fight through it and use all available resources, even if it means sacrificing something else important to me.
- 2) I would be slightly discouraged, but I would stick with it as long as I could. If I get to the midway point in the semester and I'm still struggling, I might drop one or more of the classes.

- 3) I would most likely drop one or two of the most difficult classes. If I still struggle, I would think about changing majors to something more attainable.
- 4) I would not want to be constantly struggling. I would probably give up on engineering and pursue a different career choice.

Item 3: Perseverance of Effort

Scott is not the best student, but he is hard working and can usually overcome difficulties. He is currently working with a group on a project, feels out of his element because the class is outside of his major, but the other group members are very familiar with the content. At times he feels out of place, but he doesn't want to change groups or work by himself.

If you were in Scott's position, how would you handle this?

- 1) I would either switch groups or drop the class. As much as I like the class, I can't work hard at something that causes me grief.
- 2) As much as I want to be useful in the group, I would probably end up being a burden as I find it hard to put in the needed effort when I'm not doing what I'm good at.
- 3) I typically work very hard, but since it's outside of my comfort zone it's a little harder for me to put in the extra time. I would still try though.
- 4) I am a hard worker, so I would put in the extra time to become more familiar with the content to work more effectively with the group.

Item 4: Consistency of interest

Mary has recently taken an interest in intermural sports. She enjoys it so much, that she decides to make becoming the team captain a long-term goal for her. She begins working toward becoming the team captain but realizes later that it will require a huge time commitment. After a few months, she has almost reached her goal, but realized that her interest is decreasing. She doesn't know if she should continue or quit.

If you were in her position, what would you most likely do?

- 1) I finish what I begin. Even if I find it less appealing than before, I've spent too much time to quit now. I would finish what I started without hesitation.
- 2) I would chat with my family and friends to see what they recommend. Even though I find it less appealing than before, I would most likely continue until I was team captain.
- 3) Even though I would like to finish what I started, I'm not sure it's worth it. I would stop and pursue a new interest.

- 4) Since I find it less interesting, I would quit. I would rather pursue other interests since this one is no longer appealing to me.

Item 5: Perseverance of Effort

Joe is having a very difficult time with his statistics course. His professor has given him many ways to get help, such as the assistance lab, free tutoring, and extended office hours. His professor also informed him that the class isn't required, so Joe doesn't need this class for his degree, but it could be useful later in his studies.

If you were in Joe's situation, what would you most likely do?

- 1) As soon as I found out it's something I don't need to do, I would stop. There is no point in being diligent when I don't have to be.
- 2) Since it's not required, I am more likely to give up. It's hard for me to be diligent when it's not something necessary.
- 3) Even though it's not required, I would stick with it until mid-semester, and if the extra help isn't helping I may quit.
- 4) It's going to be difficult, but with diligence I am definitely capable of succeeding. I would stick with it.

Item 6: Consistency of Interest

Diana is working on a project that she is very passionate about. Her friends even commented that she has become obsessed with it. However, after working on it for only a short time, she becomes bored of it and ultimately abandons it. This seems to happen to Diana fairly regularly that she loses interest in something she was interested in quickly.

How much do you relate to Diana?

- 1) I don't relate to her at all. When I become passionate, even obsessed with a project, I see it to the end. I don't lose interest over a short period of time.
- 2) I relate to her very little. Typically once I'm interested in something, I don't change my mind. It does happen once in a while though.
- 3) I relate to her. I find that my interests change from time to time over short periods of time.
- 4) I'm exactly like Diana. I change interests all the time, and always after a very short amount of time.

Item 7: Perseverance of Effort

Steven is in his last year of study, but is encountering a number of setbacks. It's getting to the point where he is considering delaying graduation, but if he does that it will create more problems for him. Most of the setbacks are related to his coursework.

If you were Steven, what would you most likely do?

- 1) I would deal with the setbacks in every way imaginable so I can achieve my goal of graduating this year. I don't let setbacks weigh me down.
- 2) I would try to deal with the setbacks, but I understand I may not succeed and it is better for me to postpone graduating until next semester.
- 3) I would try to achieve my goal, but I probably wouldn't be able to deal with the setbacks and I would achieve my goal at a later time.
- 4) I do not deal with setbacks very well and would most likely not achieve my goal until a much later time.

Item 8: Consistency of Interest

John has been interested in finance since his freshman year. Now that he has finished his finance introduction classes, he doesn't seem to be as interested in the field as he once was. He takes a few engineering courses and thinks that that may now be his calling. However, once he starts taking the tougher classes, he doesn't think engineering is his passion either. He simply can't stick to one academic path.

If you were John, what would you most likely do?

- 1) I would not change my major. I maintain my interests over long periods of time so I would not continue changing my majors.
- 2) I would stick with engineering a bit longer. I want to really make sure that I am not interested in it before making a change. I don't make changes lightly.
- 3) I may pursue one more engineering class to see if I really don't find it interesting anymore, but the odds are I would change majors.
- 4) I tend to change interests frequently, so I would most likely drop engineering and start focusing on something else.

Item 9: Consistency of Interest

Annabelle has always had problems with staying interested in her goals for long periods of time. Whenever she finds something that interests her, something else comes along and distracts her. Lately this has begun affecting her work because she starts projects without finishing them because she loses interests halfway through completing them.

How much do you relate to Annabelle?

- 1) I relate very much. I can't seem to stay interested in my goals for very long.
- 2) I somewhat relate to Annabelle. If I'm really passionate about something, I'm more likely to finish, but it is quite easy for me to get distracted.
- 3) I don't relate to Annabelle. If I start something, I try to finish. There may be times where I get distracted, but it's rare.
- 4) I don't relate at all. I finish everything I start, even if I lose interest over time. I must finish things I start before taking on something new.

Item 10: Perseverance of Effort

Karl has decided to add a minor to his degree plan. Several of his professors have advised him that a foreign language minor could help set him apart in the job market, so he has started studying Spanish. He is unfortunately finding it much more difficult than he had expected. It appears that he will need to put a significant amount of time into the coursework, something he had not planned on.

If you were Karl, what would you most likely do?

- 1) I am a hard worker, and I would spend as much time as needed to get the most out of the class, even if I have to sacrifice my time elsewhere. I want to get the most of out it.
- 2) I would put in the effort, but I'm not sure I would sacrifice time elsewhere. It's important to me, but as long as I get the grades I need I'm happy.
- 3) I might put in enough time to finish the minor, but if it's a lot more work than I anticipated it may not be worth it and I could drop it.
- 4) Since it's so much more of a time commitment than I first thought, I would probably change my mind on getting a minor and focus on my other courses.

Scoring Guide

Response 1 = 1 point

Response 2 = 2 points

Response 3 = 3 points

Reponses 4 = 4 points

Reverse-coded Items

2, 4, 6, 8, 10

* Response 4 from Item 1 is a bogus response. Students who select this item may not be providing valid responses to the scale.

APPENDIX D:

GRIT-S SCALE

1. New ideas and projects sometimes distract me from previous ones.

Very much like me Mostly like me Somewhat like me

Not much like me Not like me at all

2. Setbacks don't discourage me. I don't give up easily.

Very much like me Mostly like me Somewhat like me

Not much like me Not like me at all

3. I often set a goal but later choose to pursue a different one.

Very much like me Mostly like me Somewhat like me

Not much like me Not like me at all

4. I am a hard worker.

Very much like me Mostly like me Somewhat like me

Not much like me Not like me at all

5. I have difficulty maintaining my focus on projects that take more than a few months to complete.

Very much like me Mostly like me Somewhat like me

Not much like me Not like me at all

6. I finish whatever I begin.

Very much like me Mostly like me Somewhat like me

Not much like me Not like me at all

7. My interests change from year to year.

Very much like me Mostly like me Somewhat like me

Not much like me Not like me at all

8. I am diligent. I never give up.

Very much like me Mostly like me Somewhat like me

Not much like me Not like me at all

9. I have been obsessed with a certain idea or project for a short time but later lost interest.

Very much like me Mostly like me Somewhat like me

Not much like me Not like me at all

10. I have overcome setbacks to conquer an important challenge. V

Very much like me Mostly like me Somewhat like me

Not much like me Not like me at all

Reverse-coded Items

1, 3, 5, 7, and 9

APPENDIX E:

ACADEMIC MOTIVATION SCALE (AMS-C 28)

Why do you go to college?

Using the scale below, indicate to what extent each of the following items presently corresponds to one of the reasons why you go to college.

Does not Correspond at all	Corresponds a little	Corresponds moderately	Corresponds a lot	Corresponds exactly			
1	2	3	4	5	6	7	
1. Because with only a high school degree I would not find a high-paying job later on	1	2	3	4	5	6	7
2. Because I experience pleasure and satisfaction while learning new things.	1	2	3	4	5	6	7
3. Because I think that a college education will help better prepare for the career I have chosen.	1	2	3	4	5	6	7
4. For the intense feelings I experience when I am communicating my own ideas to others.	1	2	3	4	5	6	7
5. Honestly, I don't know; I really feel that I am wasting my time in school.	1	2	3	4	5	6	7
6. For the pleasure I experience while surpassing myself in my studies.	1	2	3	4	5	6	7
7. To prove to myself that I am capable of completing my college degree.	1	2	3	4	5	6	7
8. In order to obtain a more prestigious job later on.	1	2	3	4	5	6	7
9. For the pleasure I experience when I discover new things never seen before.	1	2	3	4	5	6	7

10. Because eventually it will enable me to enter the job market in a field that I like.
 1 2 3 4 5 6 7
11. For the pleasure that I experience when I read interesting authors.
 1 2 3 4 5 6 7
12. I once had good reasons for going to college; however, now I wonder whether I should
 continue.
 1 2 3 4 5 6 7
13. For the pleasure that I experience while I am surpassing myself in one of my personal
 accomplishments.
 1 2 3 4 5 6 7
14. Because of the fact when I succeed in college I feel important.
 1 2 3 4 5 6 7
15. Because I want to have “the good life” later on.
 1 2 3 4 5 6 7
16. For the pleasure that I experience in broadening my knowledge about subjects which
 appeal to me.
 1 2 3 4 5 6 7
17. Because this will help me make a better choice regarding my career orientation.
 1 2 3 4 5 6 7
18. For the pleasure I experience when I feel completely absorbed by what certain authors
 have written.
 1 2 3 4 5 6 7
19. I can’t see why I go to college and frankly, I couldn’t care less.
 1 2 3 4 5 6 7
20. For the satisfaction I feel when I am in the process of accomplishing difficult academic
 activities.
 1 2 3 4 5 6 7
21. To show myself that I am an intelligent person.
 1 2 3 4 5 6 7
22. In order to have a better salary later on.
 1 2 3 4 5 6 7

23. Because my studies allow me to continue to learn about many things that interest me.
 1 2 3 4 5 6 7
24. Because I believe that a few additional years of education will improve my competence as a worker.
 1 2 3 4 5 6 7
25. For the “high” feeling that I experience while reading about various interesting subjects.
 1 2 3 4 5 6 7
26. I don’t know; I cant understand what I am doing in school.
 1 2 3 4 5 6 7
27. Because college allows me to experience a personal satisfaction in my quest for excellence in my studies.
 1 2 3 4 5 6 7
28. Because I want to show myself that I can succeed in my studies.
 1 2 3 4 5 6 7

Key for AMS-C

- | | |
|----------------|--|
| #2, 9, 16, 23 | Intrinsic motivation – to know |
| #6, 13, 20, 27 | Intrinsic motivation – toward accomplishment |
| #4, 11, 18, 25 | Intrinsic motivation – to experience stimulation |
| #3, 10, 17, 24 | Extrinsic motivation – identified |
| #7, 14, 21, 28 | Extrinsic motivation – introjected |
| #1, 8, 15, 22 | Extrinsic motivation – external regulation |
| #5, 12, 19, 26 | Amotivation |

APPENDIX F:

NATIONAL SURVEY OF STUDENT ENGAGEMENT (NSSE)

1. During the current school year, about how often have you done the following?
Response options: Very often, Often, Sometimes, Never
 - a. Asked questions or contributed to course discussions in other ways
 - b. Prepared two or more drafts of a paper or assignment before turning it in
 - c. Come to class without completing readings or assignments
 - d. Attended an art exhibit, play, or other arts performance (dance, music, etc.)
 - e. Asked another student to help you understand course material
 - f. Explained course material to one or more students
 - g. Prepared for exams by discussing or working through course material with other students
 - h. Worked with other students on course projects or assignments
 - i. Given a course presentation

2. During the current school year, about how often have you done the following?
Response options: Very often, Often, Sometimes, Never
 - a. Combined ideas from different courses when completing assignments
 - b. Connected your learning to societal problems or issues
 - c. Included diverse perspectives (political, religious, racial/ethnic, gender, etc.) in course discussions or assignments
 - d. Examined the strengths and weaknesses of your own views on a topic or issue
 - e. Tried to better understand someone else's views by imagining how an issue looks from their perspective.
 - f. Learned something that changed the way you understand an issue or concept
 - g. Connected ideas from your courses to your prior experiences and knowledge

3. During the current school year, about how often have you done the following?
Response options: Very often, Often, Sometimes, Never
 - a. Talked about career plans with a faculty member
 - b. Worked with a faculty member on activities other than coursework (committees, student groups, etc.)
 - c. Discussed course topics, ideas, or concepts with a faculty member outside of class
 - d. Discussed your academic performance with a faculty member

4. During the current school year, about how often have you done the following?
Response options: Very often, Often, Sometimes, Never
 - a. Reached conclusions based on your own analysis of numerical information (numbers, graphs, statistics, etc.)

- b. Used numerical information to examine a real-world problem or issue (unemployment, climate change, public health, etc.)
 - c. Evaluated what others have concluded from numerical information
5. During the current school year, about how often have you done the following?
Response options: Very often, Often, Sometimes, Never
- a. Identified key information from reading assignments
 - b. Reviewed your notes after class
 - c. Summarized what you learned in class or from course materials
6. About how many hours do you spend in a typical 7-day week doing the following?
Response options: 0, 1-5, 6-10, 11-15, 16-20, 21-25, 26-30, more than 30 (hours per week)
- a. Preparing for class (studying, reading, writing, doing homework or lab work, analyzing data, rehearsing, and other academic activities)
 - b. Participating in co-curricular activities (organizations, campus publications, student government, fraternity or sorority, intercollegiate or intramural sports, etc.)
 - c. Working for pay on campus
 - d. Working for pay off campus
 - e. Doing community service or volunteer work
 - f. Relaxing and socializing (time with friends, video games, TV or videos, keeping up with friends online, etc.)
 - g. Providing care for dependents (children, parents, etc.)
 - h. Commuting to campus (driving, walking, etc.)
7. Of the time you spend preparing for class in a typical 7-day week, about how much is on an assigned reading?
Response options: Very little, Some, About half, Most, Almost all

APPENDIX G:

ACADEMIC SUCCESS QUESTIONNAIRE

The following questionnaire is designed to provide the researcher with information about yourself. Questions 1 – 3 relate to your academic career. Questions 4 – 6 relate to your demographic information. Questions 7 – 9 relate to your parents.

- 1) What is your current cumulative grade point average (GPA)?

 - 2) _____
How many classes have dropped or withdrawn from because the classes were too difficult?

 - 3) _____
How many times have you changed majors because the major was too difficult?

 - 4) _____
What is your gender association?

 - 5) _____
What is your race/ethnicity?
White
Hispanic
Black or African American
Asian or Pacific Islander
Native American or American Indian
Other
 - 6) What is your age?

18 – 24	45 - 54
25 – 34	55 - 64
35 – 44	65 or older

 - 7) What type of occupation do your parents have?
Example: Dentist, Construction Worker, Accountant
-

8) What is your parent's annual income?

9) What is the highest level of education obtained by your parents?

Did not complete high school

High school

Some College (Did not complete degree)

Associate degree (2-year degree)

Undergraduate degree (4-year degree)

Graduate degree (6+ years)

- | | |
|-----------------------------------|---|
| ___ 18. Tends to be disorganized | ___ 40. Likes to reflect, play with ideas |
| ___ 19. Worries a lot | ___ 41. Has few artistic interests |
| ___ 20. Has an active imagination | ___ 42. Likes to cooperate with others |
| ___ 21. Tends to be quiet | ___ 43. Is easily distracted |
| ___ 22. Is generally trusting | ___ 44. Is sophisticated in art, music, or literature |

Scoring:

BFI scale scoring (“R” denotes reverse-scored items):

Extraversion: 1, 6R, 11, 16, 21R, 26, 31R, 36

Agreeableness: 2R, 7, 12R, 17, 22, 27R, 32, 37R, 42

Conscientiousness: 3, 8R, 13,

18R, 23R, 28, 33, 38, 43R

Neuroticism: 4, 9R, 14, 19, 24R, 29, 34R, 39

Openness: 5, 10, 15, 20, 25, 30, 35R, 40, 41R, 44

APPENDIX I:
IRB APPROVAL

THE UNIVERSITY OF ALABAMA® | Office of the Vice President for
Research & Economic Development
Office for Research Compliance

June 11, 2018

Mitchell Porter
Dept. of ESPRMC
College of Education
Box 870231

Re: IRB#: 18-OR-211 "The Development and Validation of a Vignette-Based Academic Grit Scale"

Dear Mitchell Porter:

The University of Alabama Institutional Review Board has granted approval for your proposed research.

Your application has been given expedited approval according to 45 CFR part 46. Approval has been given under expedited review category 7 as outlined below:

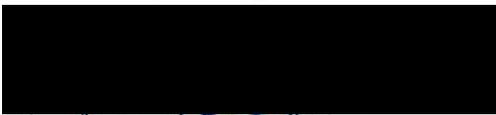
(7) Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies

Your application will expire on June 10, 2019. If your research will continue beyond this date, complete the relevant portions of the IRB Renewal Application. If you wish to modify the application, complete the Modification of an Approved Protocol Form. Changes in this study cannot be initiated without IRB approval, except when necessary to eliminate apparent immediate hazards to participants. When the study closes, complete the appropriate portions of the IRB Request for Study Closure Form.

Please use reproductions of the IRB approved stamped consent form to provide to your participants.

Should you need to submit any further correspondence regarding this proposal, please include the above application number.

Good luck with your research.



Carpanato T. Myles, MSM, CIM, CIP
Director & Research Compliance Officer

358 Rose Administration Building | Box 870127 | Tuscaloosa, AL 35487-0127
205-348-8461 | Fax 205-348-7189 | Toll Free 1-877-820-3066

CONSENT TO PARTICIPATE IN RESEARCH

You are asked to participate in a research study conducted by Mitch Porter, a PhD student from the Educational Research, Methodology, and Assessment program at The University of Alabama. Dr. Sara Tomek, Assistant Professor, is the faculty supervisor for Mr. Porter's dissertation. The results of this experiment will contribute to Mr. Porter's dissertation. You were selected as a possible participant in this study because you meet the qualifications. The qualifications are:

- 1) You are a student in an undergraduate program
- 2) You are in your junior or senior year of study

PURPOSE OF THE STUDY

The purpose of this study is to develop and validate a new scale which aims to measure a person's level of academic grit.

PROCEDURES

Each participant will be involved in a semi-structured focus group. The purpose of the focus group is to gather anecdotal evidence related to the participant's experiences and thoughts on grit. The focus group will last between 30 – 45 minutes. The focus group will be audio recorded. The experiment will take place in the Research Assistance Center in Carmichael Hall 101.

POTENTIAL RISKS AND BENEFITS

Potential discomforts and/or inconveniences include: possible boredom during the focus group, and possible anxiety due to speaking in front of others. To manage these potential discomforts you are invited to take breaks as you feel them necessary. This research will further people's understanding of grit in academia, as well as provide a better measurement tool to assess grit.

PAYMENT FOR PARTICIPATION

You will not receive any monetary compensation for participating in this study.

COSTS TO YOU

The costs to you for participating in the study is gas/travel to the university to participate in the study.

INCENTIVES

As a thank you for your time and effort for participating in this study, participants will be provided snacks and bottled water. Participants may choose to eat and drink during the focus group, or at the end of the experiment.

CONFIDENTIALITY

In order to ensure confidentiality, no names or identifying information will be recorded, and random numbers will be used instead of names for each participant.

UNIVERSITY OF ALABAMA IRB
CONSENT FORM APPROVED: 6/11/18
EXPIRATION DATE: 6/10/2019

Participants will not be permitted to review, edit or erase the audio recordings after the experiment. Only the faculty supervisor, Dr. Sara Tomek, and I will have access to the materials used in this experiment.

VOLUNTARY CONSENT

Taking part in this study is voluntary. It is your free choice. You can refuse to be in it at all. If you start the study, you can stop at any time. There will be no effect on your relations with the University of Alabama.

IDENTIFICATION OF INVESTIGATORS

If you currently have any questions or concerns about the research, or if you have any questions or concerns in the future, please ask the Principal Investigator, Mitch Porter, in person, or at mmporter2@crimson.ua.edu. You may also contact the Faculty Sponsor, Dr. Sara Tomek, at stomek@ua.edu.

RIGHTS OF RESEARCH SUBJECTS

You may withdraw your consent at any time and discontinue participation without penalty. You are not waiving any legal claims, rights or remedies because of your participation in this research study. If you have questions regarding your rights as a research subject, contact Ms. Tanta Myles, the Research Compliance Officer of the University, at 205-348-8461. You may also ask questions, make suggestions, or file complaints and concerns through the IRB Outreach website at:

http://osp.ua.edu/site/PRCO_Welcome.html

You may also email the Research Compliance office at participantoutreach@bama.ua.edu. After you participate, you are encouraged to complete the survey for research participants that is online at the outreach website or you may ask the investigator for a copy of it and mail it to the University Office for Research Compliance, Box 870127, 358 Rose Administration Building, Tuscaloosa, AL 35487-0127.

I understand the procedures and conditions of my participation described above. My questions have been answered to my satisfaction, and I freely agree to participate in this study.

Name of Subject (printed)

Signature of Subject Date

STATEMENT and SIGNATURE OF INVESTIGATOR

In my judgment the subject is voluntarily and knowingly giving informed consent and possesses the legal capacity to give informed consent to participate in this research study.

Signature of Investigator Date

UNIVERSITY OF ALABAMA IRB
CONSENT FORM APPROVED: 6/11/18
EXPIRATION DATE: 6/10/2019

CONSENT TO PARTICIPATE IN RESEARCH

You are asked to participate in a research study conducted by Mitch Porter, a PhD student from the Educational Research, Methodology, and Assessment program at The University of Alabama. Dr. Sara Tomek, Assistant Professor, is the faculty supervisor for Mr. Porter's dissertation. The results of this experiment will contribute to Mr. Porter's dissertation. You were selected as a possible participant in this study because you meet the qualifications. The qualifications are:

- 1) You are a student in an undergraduate program
- 2) You are in your junior or senior year of study

PURPOSE OF THE STUDY

The purpose of this study is to develop and validate a new scale which aims to measure a person's level of academic grit.

PROCEDURES

Each participant will complete six surveys (Grit-S, Grit-V, Academic Motivation Scale – College, National Survey for Student Engagement, the Big Five Inventory, and Academic Success Questionnaire). The participants can complete the surveys at any location that provides internet access. The participants will be provided an online link to access the surveys and complete them. The total time estimated to finish the surveys is 30 – 45 minutes.

POTENTIAL RISKS AND BENEFITS

Potential discomforts and/or inconveniences include: possible boredom while answering the survey questions, and possible anxiety due to feeling like this is a testing situation. To manage these potential discomforts, you are invited to take breaks as you feel them necessary. This research will further people's understanding of grit in academia, as well as provide a better measurement tool to assess grit.

PAYMENT FOR PARTICIPATION

You will not receive any monetary compensation for participating in this study.

COSTS TO YOU

The costs to you for participating in the study is gas/travel to the university to participate in the study.

INCENTIVES

Participants will be helping improve the line of research associated with grit, non cognitive skills, and scale development. Since participants can elect to complete the surveys in the comforts of their own home, may choose to eat and drink as they need without judgement from others.

CONFIDENTIALITY

In order to ensure confidentiality, no names or identifying information will be recorded, and random numbers will be used instead of names for each participant.

UNIVERSITY OF ALABAMA IRB
CONSENT FORM APPROVED: 6/10/18
EXPIRATION DATE: 6/10/2019

Participants will not be permitted to review, edit or erase the results of the surveys after the experiment. Only the faculty supervisor, Dr. Sara Tomek, and I will have access to the materials used in this experiment.

VOLUNTARY CONSENT

Taking part in this study is voluntary. It is your free choice. You can refuse to be in it at all. If you start the study, you can stop at any time. There will be no effect on your relations with the University of Alabama.

IDENTIFICATION OF INVESTIGATORS

If you currently have any questions or concerns about the research, or if you have any questions or concerns in the future, please ask the Principal Investigator, Mitch Porter, in person, or at mmporter2@crimson.ua.edu. You may also contact the Faculty Sponsor, Dr. Sara Tomek, at stomek@ua.edu.

RIGHTS OF RESEARCH SUBJECTS

You may withdraw your consent at any time and discontinue participation without penalty. You are not waiving any legal claims, rights or remedies because of your participation in this research study. If you have questions regarding your rights as a research subject, contact Ms. Tanta Myles, the Research Compliance Officer of the University, at 205-348-8461. You may also ask questions, make suggestions, or file complaints and concerns through the IRB Outreach website at:

http://osp.ua.edu/site/PRCO_Welcome.html

You may also email the Research Compliance office at participantoutreach@bama.ua.edu. After you participate, you are encouraged to complete the survey for research participants that is online at the outreach website or you may ask the investigator for a copy of it and mail it to the University Office for Research Compliance, Box 870127, 358 Rose Administration Building, Tuscaloosa, AL 35487-0127.

I understand the procedures and conditions of my participation described above. My questions have been answered to my satisfaction, and I freely agree to participate in this study.

Name of Subject (printed)

Signature of Subject

Date

STATEMENT and SIGNATURE OF INVESTIGATOR

In my judgment the subject is voluntarily and knowingly giving informed consent and possesses the legal capacity to give informed consent to participate in this research study.

Signature of Investigator

Date

UNIVERSITY OF ALABAMA IRB
CONSENT FORM APPROVED: 6/11/18
EXPIRATION DATE: 6/10/2019

Participate in RESEARCH!

I am looking for students to participate in a study focusing on the development and validation of a vignette-based academic grit scale. The experiment will take 30 - 45 minutes to complete.

The experiment will take place in the Research Assistance Center in Carmichael Hall 101.

Here are the **requirements**:

1. You must be enrolled in an undergraduate program
2. You must be in your junior or senior year of your program

For more information or to make an appointment to participate, please contact me, Mitch Porter, at mmporter2@crimson.ua.edu.

UA IRB Approved Document
Approval date: 6/11/18
Expiration date: 6/10/2019

Mitch Porter Academic Grit Scale Study mmporter2@crimson.ua.edu	Mitch Porter Academic Grit Scale Study mmporter2@crimson.ua.edu	Mitch Porter Mitch Porter Academic Grit Scale Study mmporter2@crimson.ua.edu	Mitch Porter Academic Grit Scale Study mmporter2@crimson.ua.edu				
---	---	--	---	---	---	---	---

Participate in RESEARCH!

I am looking for students to participate in a study focusing on the development and validation of a vignette-based academic grit scale. The experiment will take 30 - 45 minutes to complete.

The experiment will take place anywhere you want! Participants will be provided a link to complete the surveys, so you can do at your convenience in your own home!

Here are the **requirements**:

1. You must be enrolled in an undergraduate program
2. You must be in your junior or senior year of your program

For more information or to make an appointment to participate, please contact me, Mitch Porter, at mmporter2@crimson.ua.edu.

UA IRB Approved Document
 Approval date: 6/11/18
 Expiration date: 6/10/2019

Mitch Porter Academic Grit Scale Study mmporter2@crimson.ua.edu	Mitch Porter Academic Grit Scale Study mmporter2@crimson.ua.edu	Mitch Porter Mitch Porter Academic Grit Scale Study mmporter2@crimson.ua.edu	Mitch Porter Academic Grit Scale Study mmporter2@crimson.ua.edu				
---	---	--	---	---	---	---	---