

Copyright by

Eliz A. Markowitz

May 2017

PERSONALITY CHARACTERISTICS AND THE NATURE OF STANDARDIZED  
TEST SCORE INTERVENTIONS

A Dissertation Presented to the  
Faculty of the College of Education  
University of Houston

In Partial Fulfillment  
of the Requirements for the Degree

Doctor of Education

by  
Eliz A. Markowitz

May 2017

PERSONALITY CHARACTERISTICS AND THE NATURE OF STANDARDIZED  
TEST SCORE INTERVENTIONS

A Dissertation for the Degree

Doctor of Education

By

Eliz A. Markowitz

Approved by Dissertation Committee:

---

Dr. Mimi M. Lee, Co-chair

---

Dr. Susan X Day, Co-chair

---

Dr. Amber Thompson, Committee Member

---

Dr. Melissa Pierson, Committee Member

---

Dr. Robert H. McPherson,  
Dean College of Education

May 2017

This dissertation is dedicated to absent friends.

## Acknowledgements

I would first like to thank Dr. Mimi Lee for her willingness to advise me and provide support throughout my time in the doctoral program. Not only did you provide crucial support throughout my doctoral journey, but you also granted me the freedom to pursue the dissertation topic that I was most interested in.

I would also like to extend my sincere gratitude to both Dr. Amber Thompson and Dr. Melissa Pierson for initially granting me the fellowship that formed the basis for this dissertation. Furthermore, I am grateful to Dr. Thompson for giving me the liberty to craft the PPR program and pursue my interest in standardized testing over the course of my time at the university. Thank you both; this dissertation would not exist without you.

To Dr. Susan X Day, I want to express my heartfelt appreciation for your infinite encouragement, guidance, and feedback throughout my time in the program. I would not have been able to accomplish this feat without your mentorship, wisdom, and insight. You have been my champion throughout this process, and I can only hope that our friendship continues.

A special thanks goes to my mother and father. While my mother told me, “those who can, do; those who can’t, teach,” I do believe that she would approve of my progress in this field. To my father I wish to express my thanks for your infinite optimism, confidence, and love. I am truly fortunate to have you as a father and friend.

And, lastly, I would like to thank all my friends, peers, and colleagues for the encouragement and support you have provided throughout this journey. You have sustained me through the years with your honesty, friendship, and love. Each individual I have interacted with has played a role in molding me into the person I am today. For that, I am truly thankful.

PERSONALITY CHARACTERISTICS AND THE NATURE OF STANDARDIZED  
TEST SCORE INTERVENTIONS

An Abstract  
of a Dissertation Presented to the  
Faculty of the College of Education  
University of Houston

In Partial Fulfillment  
of the Requirements for the Degree

Doctor of Education

by

Eliz A. Markowitz

May 2017

## Abstract

While standardized tests (STs) intend to measure student achievement, the success of the test preparation industry calls the claim into question. The goal of test preparation is simple: for a price, provide students with content superseding techniques (CSTs) that neglect the difficult content to raise ST scores. The success of the standardized test industry indicates that high ST scores may simply reflect a personality amenable to memorization of CSTs, instead of reflecting student achievement or aptitude. This study intends to determine whether students are able to achieve higher STS with content-based or CST-based methods, and the personality traits that are associated with STS increases. The study design is a pre-post test design with three groups: a content-based group, a CST-based group, and a control group, which participated in neither intervention.  $N = 173$  participants were undergraduates in the teacher education program at the University of Houston, took equivalent pre- and post-tests consisting of PPR items from released exams, and completed two personality measures (the NEO-PI-R and the TACO inventory). There was a statistically significant difference at the  $p < 0.10$  level, set a priori, in score change for the three intervention groups:  $F(2, 124) = 2.84, p = 0.06$ . The actual difference in mean score change between the content-based intervention (0.11) and technique-based intervention (0.68) was relatively large. The effect size of the gain, calculated using Cohen's  $d$ , was 0.34, a medium effect size. Individuals are able to achieve significant STS increases without necessarily understanding the content a ST intends to assess. Eliminating STs, or developing alternative performance measures, could rectify flaws with the current evaluation system and improve learning and teaching of content.

## Table of Contents

Chapter	Page
I. Introduction.....	1
Background and Purpose of Study.....	3
Problem Statement.....	4
Significance of Study.....	5
Research Questions.....	6
Definition of Terms.....	6
Summary.....	9
II. Literature Review.....	11
Standardized Test Structure.....	12
Pedagogy and Professional Responsibilities Exam.....	12
Standardized Test Controversy.....	13
Praise.....	22
Criticism.....	15
Test-Preparation.....	19
Content-Superseding Techniques.....	20
Implications of Test Preparation.....	20
The Effect of Personality Traits on Learning.....	22
Personality Traits and Standardized Test Performance.....	23
Personality Traits and Cognitive Ability.....	25
Learnability of Personality Traits.....	26
III. Methodology.....	28
Identifying Necessary Traits for Standardized Test Score Improvement.....	29
The TACO Survey.....	31
Pilot Study.....	34
Revised TACO Survey.....	37
Final Study.....	38
IV. Results.....	42
Data Analysis.....	42
Demographic Data.....	43
NEO-PI-R, TACO, and PPR Intervention Results.....	44
Score Change by Intervention.....	45
Score Change by Personality Characteristics.....	46
V. Discussion & Conclusion.....	50
Standardized Test Score Increases by Intervention Type.....	51
Standardized Test Score Increases and Personality Variables.....	52
What do Standardized Tests Measure?.....	53
Implications of Study Results.....	55
Strengths and Limitations.....	60
Future Directions.....	62
Conclusion.....	75
References.....	77
APPENDIX A: TRAIT DISCOVERY RESEARCH & INTERVIEW PROTOCOL.....	75



APPENDIX B: PORTION OF INTERVIEW ONE.....	85
APPENDIX C: PORTION OF INTERVIEW TWO.....	98
APPENDIX D: CONSENT FORM.....	85
APPENDIX E: ORIGINAL TACO SURVEY.....	111
APPENDIX F: TACO PILOT STUDY RESULTS.....	118
APPENDIX G: PPR PRETEST/POSTTEST .....	123
APPENDIX H: REVISED TACO SURVEY.....	128

## List of Tables

Table	Page
1. Original TACO Survey .....	43
2. Average Pretest to Posttest Score Change, TACO Score, & TACO Subscale Score ..	45
3. Pearson's Correlations for Score Change Between Pre- and Post-Tests, Average TACO Subscale Scores, and Average TACO Score .....	46
4. Revised TACO Survey .....	49
5. Score Change by Intervention .....	55
6. TACO and NEO-PI-R Subscale Scores .....	55
7. Pearson's Correlations Between Personality Characteristics on TACO Survey .....	58
8. Pearson's Correlations Between TACO and NEO-PI-R Personality Subscales .....	59
9. Pearson's Correlations Between TACO and NEO-PI-R Personality Subscales and Score Change .....	60

## List of Figures

Figure	Page
1. Sample compliance question .....	42
2. Component matrix generated for factor analysis .....	47
3. Flow of subjects diagram for final study .....	52
4. Mean difference in score change by intervention type .....	57

# **Chapter I**

## **Introduction**

Standardized testing has become entrenched in all facets of education, from school-based to national exams. Developed and administered by the Educational Testing Service (ETS), the Pedagogical and Professional Responsibilities (PPR) exam is a standardized test (ST) required for all public-school educators in Texas. According to the ETS (2017a), the company that is responsible for creating and administering college admissions and college placement exams such as the SAT, graduate school admissions exams such as the Graduate Record Examination, and state-mandated exams such as the State of Texas Assessments of Academic Readiness, the content and reasoning abilities assessed on STs reflect the critical thinking skills necessary for success in a certain specialty.

The goal of a ST is to assess either student achievement or aptitude in an unbiased manner. Indeed, many institutions believe that STs accurately predict whether or not an individual is suited for a given program, or whether a student will be successful in a particular course of study (Kuncel & Hezlett, 2007). However, controversy over the validity and reliability of STs continues, as proponents of STs believe that such metrics provide valuable, objective information regarding a student's strengths and weaknesses, while critics believe that STs fail to provide meaningful insight about student knowledge (Popham, 1999; Ravitch, 2010). However, success on standardized tests may simply reflect a student's ability to memorize key words and phrases, a student personality

amenable to memorization of either content knowledge or test strategies, or a combination of these, instead of reflecting student achievement or aptitude.

While the ST industry has developed and expanded over the past century, so too has the test-preparation industry. Indeed, based on 2009 figures, the test-preparation industry is valued between \$1 billion and \$4 billion (Howard, 2014; Van Buskirk, 2009). Companies such as Kaplan, The Princeton Review, ExamCrackers, Manhattan GMAT, Khan Academy, and Testmasters have emerged as leaders of the test-preparation industry and offer myriad courses for standardized test score (STS) improvement. The goal of test-preparation companies is simple: for a price, offer courses or materials that provide students with content-superseding techniques (CSTs) that transform difficult content into manageable material.

This dissertation begins by examining the purpose, content, and structure of standardized tests (STs), and then delves into information regarding the content-superseding techniques (CSTs) promoted by the test-preparation industry to improve standardized test scores (STS). Next, current research on both the effect of personality on learning and instructor-identified personality traits necessary for STS improvement will be discussed. I will then discuss the construction and purpose of the TACO test, which assesses an individual's belief in a task's value, attitude, compliance, and open-mindedness. Then, the use of the TACO test and the prevailing five-factor personality inventory (the NEO) to predict STS improvement on the Pedagogy and Professional Responsibilities (PPR) exam will be analyzed. Finally, I will discuss the implications of these findings, including the impact of test preparation on STS improvement, the

potential benefits of adopting personality traits to improve STS, and the problems of using a ST as a measure of student aptitude.

### **Background and Purpose of Study**

Despite the fact that STs aim to assess student understanding of various content areas, the predictable nature of STs enables individuals to significantly improve standardized test scores (STS) by focusing on the format and structure of the questions. Indeed, as the success of test-preparation industry has demonstrated, the tactical skills necessary to achieve success on a ST, such as the Pedagogy and Professional Responsibilities (PPR) exam, can be acquired, and STS can improve not through a thorough understanding of the content, but through the use of CSTs.

While the format and structure of a ST can be learned, the learning style of an individual is based upon myriad factors, including personality and behavioral traits. Prior studies have shown that personality and emotion are important factors in education; in particular, an individual's traits have a significant impact on attention, which in turn affects both learning and memory (Sylwester, 1994). Indeed, numerous studies have shown that personality traits, neuroticism, extraversion, and openness to experience, play a significant role in student learning (Conard, 2006; Hofer et. al., 2012; and Nofle & Robins, 2007). Thus, it stands to reason that a student's personality traits would also play a significant role in a student's learning the content-superseding techniques (CSTs) espoused by the test-preparation industry.

The purpose of this study is twofold: (1) to determine whether individuals achieve greater STS increases using content-based or CST-based learning interventions and whether those STS increases are significant and (2) to determine whether or not specific

personality traits are linked to STS increases. If the traits that affect an individual's ability to increase STS can be identified, the role of standardized tests (STs) will be significantly different than what is usually assumed. Primarily, if an individual can magnify the identified personality traits that lead to effective use of CSTs, the potential benefits of such CSTs become greater. For example, if an individual can proactively engage in a hyper-conscientious mindset towards using CSTs on a ST, and conscientiousness toward CSTs improves STS, the student could increase her standardized test scores (STS) without improving her content knowledge (Hofer et. al., 2012).

### **Problem Statement**

Currently, students spend a significant amount of classroom time on testing, rather than higher-level learning. According to one report, students in grades K-12 spend between 60 - 110 hours per year preparing and taking standardized tests, which totals to 14 STs over 47 days of testing days per year (Nelson, 2013). However, as the success of the test-preparation industry has demonstrated, the skills necessary to achieve success on a standardized test can be learned in a shorter period of time, not through a thorough understanding of content, but through the use of CSTs.

Accordingly, the ability to learn CSTs and modify behavior to use such techniques to increase STS directly opposes the view that STs are objective measures of student achievement or aptitude. Increasing STS by simply modifying existing behavior and using CSTs would indicate that current standardized assessments are neither objective, nor are they an accurate measure of student achievement as it is commonly understood. Thus, to truly reflect students' knowledge, institutions should develop an alternative form of student assessment that validly measures student achievement based

on curricula, rather than ability to apply generalized CSTs and score well on a standardized test. A properly developed alternative form of assessment would reduce the amount of time students spend preparing for, and taking, STs, increase the amount of time students spend learning content-based curricula both in and outside of the classroom, provide more accurate information on student achievement and potential for future academic access, and encourage a more positive attitude towards schooling.

### **Significance of Study**

To date, a dearth of studies investigated the effect of the test-preparation industry on STS. While test-preparation companies tout satisfaction guarantees if score improvements are not achieved, the industry is notoriously secret regarding its processes and does not release any data regarding score improvements. Furthermore, these industries do not allow their employees or external researchers to analyze company-obtained student data or use such data for research purposes. However, the continued growth of the test-preparation industry indicates that individuals can improve STS without content knowledge. Compounding this problem are factors such as affluence and geography. Test-preparation courses are generally located in wealthy, urban centers and cost money; e.g., a SAT course in Houston, Texas, with 18 live classroom hours, costs \$899 at Kaplan, while the unlimited preparation course costs \$1599 (Kaplan, 2017b). Accordingly, students who can afford, and have access to, test-preparation courses have an advantage over other students who lack the resources necessary to participate in such courses. Such inequitable, artificially inflated scores directly oppose the goal of standardized tests, which are intended to be an equalizer, providing accurate, objective measures of student achievement regardless of student social class



Not only will this study contribute to the understanding of the role played by the test-preparation industry, but it will also provide insight into the character traits of individuals who successfully use content-superseding techniques (CSTs) to increase standardized test scores (STS), indicating that behavior modification, not knowledge acquisition, can lead to increased STS. While there have been numerous studies questioning the worth of standardized tests (STs), such studies have focused on environmental factors, rather than an individual's ability to improve STS without increasing content knowledge. Based on the results of this study, it may be necessary to revise the form of assessment used to measure student achievement, creating an assessment that is truly objective and unable to be conquered by techniques that devalue the content.

### **Research Questions**

- Do students achieve greater STS increases after a content-based or technique-based review intervention?
- Do participants benefit differentially from content- or technique-based interventions?
- Do the NEO-PI-R and TACO inventories assess similar or different characteristics?
- Are personality variables on either the NEO-PI-R or the TACO survey associated with STS increases?
- Do STS measure content mastery, mastery of CSTs, or personality variables?

### **Definition of Terms**

Content-Superseding Techniques:

Content-Superseding Techniques are those that are not taught, and would generally not be deemed appropriate, in the traditional, K-12 classroom. For the purpose of this study, content-superseding techniques are those taught in, and touted by, test-preparation companies. Often, content-superseding techniques undermine the seeming purpose of a standardized test. For example, the traditional approach to a math problem that has variables would involve the use of algebra, while the non-traditional approach to the same problem would involve the use of substitution.

#### Higher-Level Critical Thinking Skills:

Higher-level critical thinking skills include critical, logical, reflective, metacognitive, and creative thinking that are activated when individuals encounter unfamiliar problems, uncertainties, questions, or dilemmas. Successful applications of the skills result in explanations, decisions, performances, and products that are valid within the context of available knowledge and experience and that promote continued growth in these and other intellectual skills (King, Goodson, & Rohani, 1998).

#### Personality Traits:

For the purpose of this study, personality traits describe human patterns of consistency across time and situation. The Big Five personality traits are openness, conscientiousness, extraversion, agreeableness, and neuroticism (Costa & McCrae, 1992).

#### Standardized Test (ST):

A standardized test is any form of test that (1) requires all test takers to

answer the same questions, or a selection of questions from a common bank of questions, under similar administrative conditions, and (2) is scored in a “standard” or consistent manner, which makes it possible to compare the relative performance of individual students or groups of students (The Glossary of Education Reform, 2015).

#### Standardized Test Preparation Course:

A standardized-test preparation course is a course provided by a test-preparation company to an individual for a price. Unlike traditional courses, which focus on providing students with content knowledge, a standardized-test preparation course aims to provide students with standardized test-taking techniques that will improve an individual’s score on a given standardized test. A few of the most well-known test-preparation companies are Kaplan, The Princeton Review, and Testmasters.

#### Standardized Test Preparation Course Instructor:

A standardized-test preparation course instructor is an individual who teaches for a test-preparation company, and educates individuals about how to use non-traditional techniques to improve standardized test scores. Unlike educators in a traditional educational setting, test-preparation course instructors emphasize the use of test-taking techniques, rather than content knowledge.

#### Standardized Test Scores (STS):

The type of standardized test score varies from test to test. However, the two types of standardized test scores are norm-referenced and criterion-referenced. Norm-referenced scores compare a student's test performance to the performance

of a clearly defined reference group called a *norming group*. The scores of the norming group are used to devise test norms—normal, below normal, and above normal performance. The tests should have been normed on a population similar to the one taking the test. Criterion-referenced scores say something about how the person tested performed relative to an absolute performance standard determined by the test-maker. Many criterion-referenced tests would be better referred to as *content-referenced* (Time Out From Testing, 2015).

Standardized Test Score Increase:

A standardized test score increase is defined as any score increase that occurs between the initial standardized test and the final standardized test; i.e.,  $\text{Final Score} - \text{Earlier Score} = \text{Score Increase}$ .

Traditional Techniques:

Traditional techniques are those that are taught in the traditional classroom, and are generally recognized as the proper methods for approaching subject matter. For the purpose of this study, traditional techniques are those taught in the majority of K-12 educational settings, and are lauded by educators as best practices. Standardized tests are created with the intent to evaluate a student's understanding of how to implement traditional techniques. For example, the traditional approach to a math problem that has variables would involve the use of algebra, while the non-traditional approach to the same problem would involve the use of substitution.

## Summary

Despite the use of standardized tests (STs) as a measure of student achievement and aptitude, the success of the test-preparation industry demonstrates that standardized test scores (STS) can be improved through the use of content-superseding techniques (CSTs). If the personality characteristics that affect the effectiveness of using CSTs on STs can be identified, and individuals can proactively make use of traits that are linked to STS improvement, the way in which STS are interpreted will be altered significantly.

## **Chapter II**

### **Literature Review**

Standardized testing is one of the most pervasive, and controversial, aspects of the public education system in the United States. While The Department of Education (DOE) created the National Assessment of Educational Progress test in 1969 to identify the educational accomplishments of students in the public institutions (Kumeh, 2011), the No Child Left Behind Act (NCLB) of 2002 increased the DOE's authority in public K-12 education. As part of NCLB, "states are required to test students annually and demonstrate continual progress toward a Federal goal of all students reaching 'proficiency' on state-level exams by 2014" (Hickok & Ladner, 2007, p. 65). Since NCLB was passed, in an apparent effort to make educational institutions and students more accountable, standardized testing has become a cornerstone of public education in America. Despite the prevalence of standardized tests (STs) in all facets of education, enormous controversy exists surrounding their use as a valid and reliable measure of student aptitude or achievement.

Compounding this controversy is the advent of the test-preparation industry, which provides individuals with materials and courses for overcoming the hurdles faced on STs. While many consider strategic item rules to be an effective means of improving standardized test scores (STS), the implications of such courses are numerous. The ability to increase STS with strategic item analysis indicates that ST questions can be conquered with test-taking techniques and content-superseding techniques (CSTs), thus negating the validity of a ST as an objective measure of student aptitude or achievement. Presently,

there is no research regarding the skills necessary for successfully learning the techniques promoted by the test-preparation industry. However, prior studies have indicated that an individual's personality traits have a significant impact on academic achievement, subject-specific test scores, and the way in which learning occurs.

### **Standardized Test Structure**

The average standardized test (ST) is composed of multiple-choice questions and administered under regulated conditions with the intent of providing both a valid and reliable metric for peer-to-peer comparison (FairTest, 2017; Popham, 1999). STs are either norm-referenced or criterion-referenced; norm-referenced tests compare a student's performance to that of peers while criterion-referenced tests gauge student performance by comparing responses to a predetermined set of criteria (Bond, 1996; FairTest, 2017) that define competence in the test's subject area.

### **Pedagogy and Professional Responsibilities Exam**

The Texas Education Agency requires that individuals who teach in Texas public schools take, and pass, the Pedagogy and Professional Responsibilities (PPR) exam. In order to register for the PPR exam, students must be enrolled in, or have completed, an educator preparation program. A teaching program's accreditation is partly based on its student's passage of the PPR. The content of the PPR exam reflects this requirement; questions on the PPR exam are related to curriculum design, instruction implementation, classroom environment, and teacher roles and responsibilities. The structure and format of the PPR exam is straightforward: the computer-administered test requires individuals to answer 100 multiple-choice questions, 10 of which are experimental questions used in future test development that do not affect the overall score, over a five-hour period.

## **Standardized Test Controversy**

While highly used, the use of standardized test scores (STS) as a measure of student achievement or aptitude remains a controversial issue among educators. While proponents and critics of standardized tests (STs) offer a plethora of reasons to support their respective sides, the crux of the issue is the use of a ST as a valid metric for gauging student knowledge and ability.

**Praise.** The most recent push for standardized testing began with the enactment of NCLB, which promised sweeping educational reform and an improvement in the quality of public education. As Rod Paige, representing the U.S. Department of Education, stated in his 2003 letter to parents, the focus of NCLB is “to see every child in America—regardless of ethnicity, income, or background—achieve high standards.” Furthermore, NCLB, and its successor Race to the Top (RTTT), holds institutions and educators accountable for student performance and annual yearly progress. As part of the mandate, progress and proficiency in subject matter is assessed annually for all students in grades K-12 via STs concentrated on both reading and mathematics. If an institution demonstrates adequate yearly progress among students, the federal government grants institutions flexibility in how federal funds are used. Proponents of STs champion NCLB and the associated yearly exams, stating that the use of the STs holds schools accountable for student learning and providing high quality education. Indeed, Hanushek and Raymond (2005) analyzed state achievement growth, as measured by the National Assessment of Education progress, and found that accountability systems had a clear positive effect on student achievement-despite their failure to account for unintended



outcomes such as higher exclusion rates from standardized testing and increased drop-out rates.

In response to critics who protest school curricula focusing on test content instead of subject matter, proponents of STs believe that skills reflected by test scores are necessary for growth. Indeed, advocates of NCLB also argue that the content and structure of STs reinforces essential skills necessary for success and should be the focus of classroom education. The Center for Public Education (2006) dismisses concerns over narrowed curricula and teaching to the test, noting that STs focus education on essential skills by eliminating classroom activities that do not yield measurable gains, and motivate educators and students to put forth more effort. Indeed, proponents of using STS for accountability purposes cite the dramatic increase in state test scores since the enactment of NCLB as evidence of increased student learning (Center on Education Policy, 2008, 2009).

When asked about the objectivity of STs, proponents laud the objective and unbiased nature of STs, stating that the machine-graded, multiple-choice format eliminates bias and subjectivity often present in the scoring of non-standardized tests. According to the Glossary of Education Reform (2015), “a standardized test is any form of test that (1) requires all test takers to answer the same questions, or a selection of questions from common bank of questions, in the same way, and that (2) is scored in a “standard” or consistent manner, which makes it possible to compare the relative performance of individual students or groups of students”. Unlike many tests encountered in educational settings, STs are composed by skilled psychometricians, tested, refined, and retested in an attempt to ensure questions are both fair and objective (Meador, 2016;

Phelps, 2002). Supporters of STs cite the objective nature of such exams, which should provide an equal grounding for students, regardless of ethnicity, socioeconomic status, or geographic location. When questioned about the use of STs for minorities and special needs students, former Chancellor of the Washington D.C. public school system, Michelle Rhee (2011), argues that providing alternative tests would create two, unequal systems: separate is not equal, and creating specialized tests would be discriminatory.

Institutions often use STs as a metric of student aptitude or achievement because they are cost-effective to both design and score (Garland, 2012). After all, STs require little humanpower to score, as a human is only needed to ensure student responses are machine-scanned. In order to properly score a subjective, short- or long-response test, it would be necessary to recruit, train, and hire individuals to score with consistency—an expense many institutions would rather not incur.

**Criticism.** In theory, standardized test scores (STS) provide parents, educators, and the government with information regarding a student's skills relative to his/her peers. While a standardized test (ST) is intended to identify a student's strengths and weaknesses, and provide educators with information for effective change, the nature of such tests often fails to provide meaningful insight about student knowledge (Kastenbaum, 2012; Koretz, 2002, 2005; Popham, 1999). In addition to the questionable validity of STs, such tests are also inherently biased against individuals of a lower socioeconomic status, reward rote memorization instead of higher-level critical thinking skills, and dilute the quality of instruction (Dexheimer, 2009; Ravitch, 2010; Strauss, 2011).

Since the enactment of NCLB, and to the detriment of all involved, ST preparation has become saturated into all aspects of public school curricula. In an effort

to achieve adequate yearly progress and reach the 2014 proficiency goal, educators have been concentrating on preparing students for STs rather than focusing on increasing student knowledge across a breadth of subject matter. Furthermore, while the results of STs are used to assess the progress of students, they are also used to evaluate the efficacy and worth of both educators and institutions—despite agreement among economists, statisticians, and psychometricians that standardized test scores (STS) are not a valid gauge of teacher effectiveness (Akers, 2016; Baker, et al., 2010; Hull, 2013).

Furthermore, even when U.S. policymakers have offered incentives to students, teachers, and schools to increase STS, these incentive systems have had either small effects or no effect on student learning (Hout, Elliot, & Freuh, 2012). Accordingly, educators and administrators are placing increasing importance on ST in the classroom since their jobs and pay often depend on student performance (Dexheimer, 2009). Unsurprisingly, the goal of 100% proficiency in 2014 was not met, and many critics claim that the quality of instruction has experienced a sharp decline (Akers, 2016; Ravitch, 2010). Rather than providing students with an enriching educational experience, educators have begun to focus their efforts on teaching for STs. Indeed, a Policy Research Brief produced by the National Council of Teachers of English (2014) observes that research shows the effects of high-stakes standardized testing include, “changing the nature of teaching, narrowing the curriculum, and limiting student learning.” In the past, educational institutions were known to provide individuals with a breadth of skills and knowledge associated with higher-order thinking and complex thought. Now, however, educators are “teaching to the test,” that is, excessively drilling students over ST material, minimizing content knowledge, and promoting rote memorization as a valuable form of learning (Crocco &

Costigan, 2007; Jennings & Bearak, 2014; McNeil, 2002; Nelson, 2013). According to a 2007 study by the Center of Education Policy, 44% of school districts have eliminated two and a half hours of science, arts, and social studies instruction each week in order to focus on math and reading questions that will occur on ST (Ravitch, 2010). Thus, in addition to having learning measured by standardized tests (STs), students are receiving an education in which both creative and higher-level cognitive thinking skills are downplayed, rather than promoted.

Intended to be an objective measure of student skills or knowledge, STs are often used to gauge student performance against peers. However, due to both the lack of a standardized national curriculum, the diversity of student populations throughout public schools in the United States, and the range of psychometricians constructing ST questions, a ST cannot be considered objective. The lack of an identical, national curriculum indicates that psychometricians are creating STs based upon a generalized subject matter and populations, rather than the actual structure existing within the range of public schools (Koretz, 2013). Constructing a test based on a generalized idea of subject matter, rather than reflecting individual students and schools, has been shown to penalize impoverished, racially diverse, and special education students (Strauss, 2011). Due to the diverse population of the United States, melding a one-size-fits-all standardized test (ST) into all schools is both impractical and discriminatory toward minorities. Furthermore, the bias that exists within the content and administration of a ST exacerbates the problem of evaluating a student's skills against those of peers. Daniel Koretz (2002) notes that it is possible to obtain non-comparable scores due to differences in test administration, date of administration, or other factors. Thus, if the testing

conditions themselves are not standardized, the results are invalid for peer-to-peer comparison and should not be used as a metric of student ability. Administrators can further devalue the validity of STs by skewing the results or lowering the standards for proficiency. In one disturbing case, a review of 200 elementary schools in Dallas, Texas, shows that even the school lowest on the fourth-grade TAKS mathematics exam achieved a 93% passing rate. The caveat? Students only needed to answer fewer than half the questions on the exam correctly in order to achieve a passing score (Hickok & Ladner, 2007). On the surface, many schools appear to be improving the quality of education. In reality, however, lowering the criterion for proficiency to meet annual yearly progress goals undermines the purpose of NCLB and tarnishes the field of education.

Critics of STs also cite norm-referenced testing as an inadequate measure of student performance. While norm-referenced testing may be useful for determining information about students within a particular institution in comparison to the nation's students as a whole, "the validity of the score . . . depends on whether or not the content of the norm-referenced test matches the knowledge and skills expected of the students in a particular school system" (Bond, 1996, p. 7). As FairTest-National Center for Fair & Open Testing (2017) notes, there are numerous issues with the use of norm-referenced STs as a measure of student achievement, including that norm-referenced tests focus too heavily on routine procedures and memorization, lower academic expectations, and cause teachers to de-emphasize thinking and application of knowledge. Thus, if teachers in a particular system refuse to teach to the nationally normed test, they and their students are penalized. Since individuals from across the nation are required to take identical STs,

despite receiving non-standardized education, the use of STS as reflections of learning is flawed.

### **Test Preparation**

At this time, a dearth of literature measures the effectiveness of test-preparation courses and materials on standardized test score improvement. The general mission of test-preparation companies is to help students achieve high scores on standardized tests (STs) through test-specific courses and materials. Instead of honing academic knowledge, test-preparation courses emphasize content-superseding techniques that transform seemingly complex questions into more simplistic ones; for example, using the strategy of eliminating all answers with a certain key word like “always,” instead of thinking about the question’s substance.

In order to provide students with quality material that effectively improves student scores, test-preparation companies evaluate and dissect ST questions and, in turn, develop techniques for overcoming challenging content. Techniques provided by test-preparation companies are highly successful as STs only cover a limited amount of subject matter in a perfunctory manner and, therefore, measure low-level cognitive processes (Brady, 2011) such as memorizing simple definitions of a pool of common terms.

Unfortunately, test-preparation courses are not universally available, with attendance limited by both price and geography. Indeed, for a price, students can enroll in an extra-curricular course that aims to improve STS. So successful in recruiting consumers are test-preparation companies that, as of 2009, the test-preparation industry is valued between \$1 and \$4 billion annually (Brody, 2009).

### **Content-Superseding Techniques**

Despite the content-based nature of the Pedagogy and Professional Responsibilities (PPR) exam, the test-preparation industry has developed techniques for tackling such reasoning questions on standardized tests (STs). On the PPR, individuals are required to answer multiple-choice questions regarding information contained within the provided question stems and informative stimuli. The length of both PPR question and stimuli is limited and, therefore, only a brief overview of a subject is provided.

One straightforward technique promoted by test-preparation companies revolves around *automatic wrong answers*, or words that indicate answer choices are automatically wrong. For example, ironically, an answer choice that suggests that a teacher's choice of test should have *a multiple-choice assessment* will be incorrect on the PPR, because multiple-choice assessments evaluate lower-level critical thinking skills and, ideally, teachers want to assess a student's higher-level critical thinking skills. Similarly, any answer choice on the PPR that suggests the use of *worksheets* or *handouts* will be incorrect. By memorizing and identifying these automatic wrong answers, individuals can correctly answer questions on the PPR while barely reading the question itself.

### **Implications of Test Preparation**

The implications of such content-superseding techniques (CSTs) are both troublesome and far-reaching. By using test-taking strategies, students have the ability to score highly without comprehending the higher-order concepts of the subject matter. Consider another standardized test, the SAT, which is used by colleges and universities for admission purposes. The quantitative section of the SAT consists of 54 questions, of

which, on average, 17 questions are intended to evaluate a student's knowledge of algebra. Based on data from Testmasters (2017), if a student can effectively use the substitution technique on eight additional questions, the math score alone could improve by upwards of 150 points, on a 200-800 point scale. Accordingly, it is entirely possible that a student who employs multiple CSTs could improve his/her overall SAT score by a tremendous amount without understanding the skills the SAT aims to assess.

So effective are test-preparation companies at helping students increase standardized test scores (STS) that the majority of these companies offer students some form of guarantee if they attend all classes, complete all homework and diagnostic exams, and do not attain higher STS on the real exam. Consider the SAT. For this exam, both Kaplan (2017b) and The Princeton Review (2017b) offer a money-back guarantee if students do not score higher on the official test, while Testmasters (2017) guarantees that students will achieve a 200-point increase. The willingness of test-preparation companies to guarantee improvement indicates that standardized-test preparation courses are effective in improving a student's STS.

One related, noteworthy, implication involves the inequity of STs and test-preparation companies; courses at test-preparation companies are costly and are geared toward affluent populations rather than those with a low socio-economic status. As of Spring 2017, an 18-hour SAT Classroom Course costs \$899 at Kaplan Test Preparation (2017), a 30 hour SAT Classroom Course at The Princeton Review (2017b) costs \$1099, and a 33-hour SAT course at Testmasters (2017b) costs \$749. Despite the fact that many students may wish to partake in a standardized test preparation course in order to improve the chances of acceptance at a top-tier university, the financial burden associated with



such courses can make such desires impossible. Accordingly, students who partake in such courses have the benefit of instruction on CSTs, while students who are unable to afford, and learn from, such courses do not.

### **The Effect of Personality Traits on Learning**

Prior research on the links between intelligence, personality, knowledge, and interest has shown that individual differences in personality may influence academic performance (Conard, 2006; Chamorro-Premuzic & Furnham, 2003; Nofle & Robins, 2007; Hofer et al., 2012). While much research has been done regarding the relationship between standardized test scores (STS) and intelligence, little research has been done regarding personality traits that affect either STS or one's ability to improve STS. However, over the past couple of decades, researchers have studied the effect of personality and behavioral traits on various relevant aspects of student performance.

Research has established that personality, the predictive power of which has little to do with intelligence or other aspects of cognitive ability, has significant influence on success at both work and school. One of the most oft used personality inventories is the NEO Personality Inventory-Revised, which measures the Big Five personality traits—openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (Costa & McCrae, 1992). Many studies have used the NEO Personality Inventory-Revised to determine how the Big Five personality traits can predict various aspects of academic performance.

### **Personality Traits and Academic Performance**

One of the initial studies regarding the impact of the Big Five personality traits and their relation to academic performance was conducted by Chamorro-Premuzic and

Furnham (2003), who hypothesized the following: (1) Neuroticism will be negatively and significantly related to academic performance; (2) Extraversion will be negatively and significantly related to academic performance as measured by written examination; (3) Openness to experience will be positively and significantly related to academic performance; (4) Agreeableness will not be significantly related to academic performance; (5) Conscientiousness will be positively and significantly related to academic performance; and (6) The Big Five trait profile will significantly predict academic performance.

In order to test these hypotheses, academic performance was measured by five three-hour exams, and personality data was collected using the NEO Personality Inventory-Revised. The results were (1) Neuroticism was significantly and positively correlated with academic performance; (2) Extraversion was *only* significantly, negatively correlated with exam marks; (3) Openness was not significantly correlated with academic performance; (4) Agreeableness was not significantly correlated with exam grades; (5) Conscientiousness was moderately, positively, and significantly related to academic performance; and (6) Personality profile (the pattern of five traits together) accounted for 13% of the variance in overall totaled exam results.

The outcome of the study indicates that personality characteristics can be useful predictors of academic success. In particular, the study gives credence to the idea that personality profiles are predictors of both exam outcomes and academic performance. Of particular interest are the traits of Neuroticism and Conscientiousness, which appear to play a significant role in academic performance.

### **Personality Traits and Standardized Test Performance**

Studies such as Chamorro-Premuzic and Furnham's (2003) have indicated a significant correlation between personality traits and academic achievement, including GPA, exam results, and course performance. However, such studies did not relate personality traits to STS. Indeed, while there are few studies that analyze the relationship between personality traits and their impact on STS, some studies do incorporate the outcome of STs as part of their analysis.

One such study, conducted by Conard (2006), investigated the incremental validity of Big Five traits for college GPA, course performance, and class attendance over academic ability, as measured by the SAT. Conard's study also tested the behavioral mediator of attendance between personality characteristics and academic performance in undergraduate psychology classes. Subjects took the NEO Personality Inventory to measure the Big Five traits, course attendance was recorded, and self-reported high school GPA and SAT scores were used to measure academic ability. The results were as follows: (1) Conscientiousness and SAT had direct effects on GPA, and attendance partly mediated the relationship between conscientiousness and GPA; (2) SAT scores and Conscientiousness had direct effects on course performance, with attendance mediating the relationship; and (3) Significant relationships were found between Conscientiousness and attendance and between Conscientiousness and course performance. One of the greatest implications of this study relates to the finding that Conscientiousness operates through behavior, like course attendance.

Similar findings were found in Nofle and Robins' (2007) study, which examined the ability of the Big Five personality traits to predict academic outcomes. Specifically, their study aimed to explore the impact of the Big Five on both SAT scores and GPA. Four

studies were conducted, in which participants were required to take the NEO Personality Inventory, and self-report high school GPA, current college GPA, and SAT scores. The results of the study showed that (1) Only Openness was consistently related to higher SAT verbal scores, and there was a weak tendency for individuals with low Extraversion and Agreeableness to score higher on SAT Verbal; (2) No personality trait was related to SAT math scores; and (3) Conscientiousness was a slightly stronger predictor of GPA than SAT scores.

The implications of these studies are numerous. Primarily, the fact that certain personality traits were found to effectively predict GPA and SAT scores indicates a definite interplay between personality traits and measures of achievement. While the studies applied different variables to operationalize academic performance, the studies showed a significant correlation between the personality traits of Conscientiousness and Openness and improved STS. If personality traits have a significant impact on the way in which an individual learns information and performs academically, it is possible that personality traits also have a significant impact on an individual's performance in a ST preparation course. Rather than using academic knowledge, students can achieve large score increases on STS using the CSTs learned outside of the classroom if they are able to employ certain personality traits.

### **Personality Traits and Cognitive Ability**

The results of the aforementioned studies were further confirmed by Hofer et al.'s (2012) study, which aimed to evaluate the predictive power of cognitive ability, self-control, use of time, academic procrastination, and motivational interference, for both self-reported grades and STS. Each student took a test of cognitive ability, a measure of self-control strength, and three measures that assessed the use of time structure, academic

procrastination, and experience of motivational interference. In addition, the subjects were given a version of the Third International Mathematics and Science Study (TIMSS) standardized test at both the beginning and end of the school year.

The results of the study showed that personality variables were better predictors of grades than of TIMSS scores, while cognitive variables were better predictors of TIMSS scores than of grades. Furthermore, the results confirmed the hypotheses that cognitive ability is a better predictor of STS than self-reported grades, that self-control strength is a better predictor of self-reported grades than of STS, and that all the aforementioned variables predict self-reported grades. Perhaps most interesting, however, was the connection the researchers made between cognitive ability and self-control strength. The researchers believe that cognitive ability can be trained effectively through educational interventions aimed at self-control.

In Hofer et al.'s (2012) study, students were given multiple measures focused on determining the strength of personality variables for each student. However, the administered measures do not reflect personality characteristics as labeled in the NEO Personality Inventory. Accordingly, it is possible that personality characteristics could have a potential effect on achieving STS increases through test preparation. Indeed, if students bring certain personality traits to bear, learning the CSTs espoused in ST preparation courses could allow students to increase STS (Klaeur & Phye, 2008).

### **Learnability of Personality Traits**

Undoubtedly, personality characteristics have an effect on STS, GPA, and academic performance. Based on the aforementioned studies, it is clear that certain traits play a significant role in the outcome of academic performance, including STS. In these

studies, Conscientiousness, Agreeableness, and Openness appeared to have the most substantial impact on increased STS. One of the greatest implications of these studies relates to the finding that Conscientiousness plays a major role in all forms of academic performance and operates through behavior.

As noted in both Klaur and Phye (2008) and Hofer et al.'s (2012) studies, it may be possible for students to train their cognitive abilities. If students can effectively employ specific personality characteristics, they can, in turn, increase STS by adopting CSTs instead of furthering content knowledge, thus negating the use of a ST as a measure of a student's academic achievement or aptitude.

## **Chapter III**

### **Methodology**

The purpose of this study was two-fold: (1) to determine whether individuals achieve greater STS increases using content-based or content superseding technique (CST)-based learning interventions and whether those STS increases are significant and (2) to determine whether or not specific personality traits are linked to STS increases in either content-based or CST-based interventions. Accordingly, the research questions explored in this study were

1. Do students achieve greater STS increases after a content- or technique-based intervention preparing for the PPR?
2. Do participants benefit differentially on PPR scores after content- or technique-based interventions?
3. Do the NEO-PI-R and TACO inventories assess similar or different characteristics?
4. Are personality variables on either the NEO-PI-R or the TACO survey associated with STS increases on the PPR?

Using the information gleaned from the aforementioned research questions, the ultimate question of the study is to determine whether STs measure content mastery, mastery of CSTs, or personality variables.

The following chapter will outline the methodology used in the study, including the (1) process used for identifying traits present in individuals who obtain score increases after participating in test preparation; (2) the development and purpose of the TACO survey, a measure intended to predict an individual's ability to increase test scores

using test-preparation techniques; (3) the pilot study; (4) the revised TACO survey; (5) and (6) the final study using both the NEO-PI-R and TACO surveys to determine if personality factors predict ability to increase STS with CSTs.

It should be noted that I have worked for a test-preparation company for over a decade, teaching, training future instructors, and working in research and development. Accordingly, I have unique insight into the methods and tools developed and implemented by test-preparation companies in their materials and courses. In combination with my academic background, most significantly my degree in Curriculum and Instruction – Learning, Design, and Technology, I am able to teach courses that focus either on content or on CSTs. For the purposes of this study, I conducted all interviews and taught all courses. This choice almost eliminated any experimenter effect, since I am practiced at separating the two types of intervention.

### **Identifying Necessary Traits for Standardized Test Score Improvement**

In order to determine the characteristics that are present in students who achieve large score increases using test-preparation methods, I interviewed a number of individuals who work at test-preparation companies. For the purposes of this study, we believed that the ability to glean insight from individuals that goes beyond statistical data allows for a more descriptive, thorough analysis than would be captured by any existing measure. Furthermore, these insights provided a real-world basis for creating items in a quantitative measure specifically designed for participants in test preparation interventions.

In a preliminary research project, we were interested in determining instructor-identified personality traits that allow students to succeed using material taught in test-



preparation courses. This portion of our study used the framework of critical qualitative research outlined by Carspecken (1996) in his text, *Critical Ethnography in Educational Research*. In brief, this methodology allows a researcher to explore how individuals subjectively view a specific concept (a core tenet of qualitative research), and from their discourse to glean themes or trends. Information regarding participants and other design aspects of this project can be found in Appendix A. The interview protocol (Appendix A) used was designed based on Carspecken's (1996) methodology, which states that an interview protocol should consist of domains, or overarching topics of interest. For this study, the major domains included instructional practices, the purpose and structure of a test-preparation course, a comparison of preparation methods, and personality traits, or characteristics, present in students who experience large score increases with standardized test preparation. The interview protocol was elaborated throughout every interview (portions of which can be seen in Appendices B and C) in an attempt to follow the thoughts of individual interviewees.

Four common themes arose regarding characteristics present in students who achieve the greatest increases in standardized-test scores (STSs). All of the interviewed instructors felt that students who receive the greatest STS increases (1) have high belief in the value of standardized-test preparation, (2) maintain positive attitude toward material and the instructor, (3) are compliant and use the content-superseding techniques (CSTs), regardless of previous familiarity with the technique, and (4) possess an open mind and willingness to try alternative techniques. Based on these four themes, we developed a survey that measures the strength of these characteristics in individuals.

The following definitions of Task Value, Attitude, Compliance, and Open-

Mindedness were determined after analyzing semi-structured interviews with test-preparation instructors and distilling the most important qualities for student score improvement. Task Value is high when a student does the following: (1) appreciates the worth of the task, regardless of whether or not it is interesting; and (2) acknowledges that understanding the task will serve a greater purpose or aid in future work. Attitude is high when a student does the following: (1) believes that the task is feasible and (2) approaches the task with a sense of confidence. Compliance is high when a student does the following: (1) adheres to task requirements, regardless of familiarity with method; (2) uses new methods to complete a task, regardless of comfort with method; and (3) strives to master new methods for task completion. Open-mindedness is high when a student does the following: (1) demonstrates willingness to attempt new methods; (2) analyzes whether or not a task can be solved in alternative ways; and (3) expresses a desire to expand knowledge.

### **The Original TACO Survey**

Based on the instructor surveys, I developed the TACO survey, shown in Table 1 and Appendix E, to gauge the strength of an individual's character in four areas: Task Value, Attitude, Compliance, and Open-Mindedness. The original TACO survey

I make sure I keep up with the weekly readings and assignments for courses.

Not at all like me
Not like me
Not much like me
Neutral
Somewhat like me
Like me
Just like me

*Figure 1. Sample compliance question.*

consisted of four demographic questions and 40 statements that required individuals to rate their behaviors and attitudes on a scale that ranges from *Not At All Like Me* to *Just Like Me*, as shown in Figure 1; student responses were later converted to numerical scores, where *Just Like Me* = 7, *Neutral* = 4, and *Not at All Like Me* = 1. The questions that required individuals to rate their behaviors and attitudes were adapted from the Motivated Strategies for Learning Questionnaire (Pintrich, 1990), which has been used to determine student study habits, learning skills, and motivation for class work. Each characteristic (Task Value, Attitude, Compliance, and Open-Mindedness) was measured with a series of ten questions, all of which were scored in the same direction, and, based upon student responses to the four sets of ten questions, both subscale scores for each personality characteristic and a total TACO score was determined.

Table 1

*Original TACO Survey*

Subscale	TACO Survey Questions
<b>Task Value</b>	<ol style="list-style-type: none"> <li>1) In class, I prefer course material that really challenges me so I can learn new things.</li> <li>2) In class, I prefer course material that arouses my curiosity, even if it is difficult to learn.</li> <li>3) I think I will be able to use what I learn in one course in other courses.</li> <li>4) It is important for me to learn the course material in class.</li> <li>5) I think the course material in this class is useful for me to learn.</li> <li>6) Understanding the subject matter of this course is very important to me.</li> <li>7) I ask myself questions to make sure I understand the material I have been studying in this class.</li> <li>8) I ask the instructor to clarify concepts I don't understand well.</li> <li>9) I make good use of my study time for this course.</li> <li>10) I usually study in a place where I can concentrate on my course work.</li> </ol>
<b>Attitude</b>	<ol style="list-style-type: none"> <li>1) I am very interested in the content area of this course.</li> <li>2) If I study in appropriate ways, then I will be able to learn the material in this course.</li> <li>3) It is my own fault if I don't learn the material in this course.</li> </ol>

- 
- 4) If I try hard enough, then I will understand the course material.
  - 5) If I don't understand the course material, it is because I didn't try hard enough.
  - 6) I'm certain I can understand the most difficult material presented in the readings for this course.
  - 7) I'm confident I can understand the basic concepts taught in this course.
  - 8) I'm confident I can do an excellent job on the assignments and tests in this course.
  - 9) I'm certain I can master the skills being taught in this class.
  - 10) When I study for this class, I set goals for myself in order to direct my activities in each study period.

### **Compliance**

- 1) When I study for this class, I practice saying the material to myself over and over.
- 2) When studying for this class, I read my class notes and the course readings over and over again.
- 3) I memorize key words to remind me of important concepts in this class.
- 4) I make lists of important terms for this course and memorize the lists.
- 5) If I don't understand the course material, it is because I didn't try hard enough.
- 6) When I become confused about something I'm reading for this class, I go back and try to figure it out.
- 7) If I get confused taking notes in class, I make sure I sort it out afterwards.
- 8) I make sure I keep up with the weekly readings and assignments for this course.
- 9) I attend class regularly.
- 10) Even when course materials are dull and uninteresting, I manage to keep working until I finish.

### **Open Mindedness**

- 1) I often find myself questioning things I hear or read in this course to decide if I find them convincing.
  - 2) When a theory, interpretation, or conclusion is presented in class or in the readings, I try to decide if there is good supporting evidence.
  - 3) Whenever I read or hear an assertion or conclusion in this class, I think about possible alternatives.
  - 4) I try to play around with ideas of my own related to what I am learning in this course.
  - 5) I try to relate ideas in this subject to those in other courses whenever possible.
  - 6) In a class like this, I prefer course material that really challenges me so I can learn new things.
  - 7) In a class like this, I prefer course material that arouses my curiosity, even if it is difficult to learn.
  - 8) I always try to understand what the teacher is saying, even if it doesn't make sense.
  - 9) I prefer knowing many approaches to material than knowing a single approach.
  - 10) I generally find that there are multiple ways to approach a given task.
- 

### **Pilot Study**

For the purposes of the scale development study, a convenience sample was used, including undergraduate education students at the University of Houston (UH) who are required to attend a CST-based Pedagogy and Professional Responsibilities (PPR) review session. Pre- and post-tests, each consisting of ten PPR questions, were developed to determine individual score improvement following standardized test preparation.

Both the PPR pre-test and the TACO survey were administered before the PPR intervention; students were given 20 minutes to complete the TACO survey and 20 minutes to complete the ten-question PPR pre-test. Following the intervention, students were given 20 minutes to complete the ten-question PPR post-test. In an attempt to ensure accurate data collection and minimize administration error, the pre-test, post-test, and TACO survey were all administered with Qualtrics, an online survey tool that allowed students to complete all three surveys with any internet-connected device.

While 113 participants completed the TACO survey, only 54 completed the pre-test, post-test, and TACO survey. Of those who completed all three surveys, 87% were female and 13% were male. There were a number of ethnicities represented, with 42.6% Hispanic/Latino, 33.3% Caucasian, 9.3% African American, 9.3% Asian, and 5.6% Other. For each participant, the average score was computed across all four subscales of Task Value, Attitude, Compliance, and Open-Mindedness. Likewise, each participant's TACO score is simply the average score of all 40 questions. Score change was computed by finding the change between the pre- and post-test.

Table 2

*Average Pre- to Post-Test Score Change, TACO Score, and TACO Subscale Scores.*

Activity	<i>M</i>	<i>SD</i>
Score Change	-.63	2.309
TACO	5.84	.563
Task Value	5.94	.658
Attitude	5.84	.686
Compliance	5.78	.691
Open-Mindedness	5.81	.725

*Note.* *N* = 54 for all analyses.

Table 2 shows the average score change between the pre- and post-test was -0.63, which could be due to a number of reasons. Primarily, while items in both the pre- and post-test had equivalent content, the level of question difficulty between the pre- and posttests were not evaluated. Without including either reverse scaled items within the TACO survey or establishing equivalent form reliability in the pre- and post-PPR tests, the relationship between the TACO scores and score change remains unclear.

After determining the average scores for each participant, correlations between the overall TACO score, the subscales, and the score change were found. As shown Table 3, and the graphs in Appendix F, there are slight, but non-significant, positive correlations between score change and both Task Value and Compliance, and significant negative correlations between score change and Attitude, Open-Mindedness, and the overall TACO score. Furthermore, the correlations between the subscales generally range

from 0.3-0.6, indicating that the items are related, but not overlapping or overly disparate (Tabachnick & Fidell, 1996).

Table 3

*Pearson's Correlations for Score Change Between Pre- and Post-Tests, Average TACO Subscale Scores, and Average TACO Score*

	Personality Characteristics				
	Task Value	Attitude	Compliance	Open-Mindedness	TACO Average
Score Change	.073 (.299)	.225 (.051)	-.137 (.161)	.210 (.063)	.116 (.203)
Task Value	1.000 .	.635** (.000)	.614** (.000)	.535** (.000)	.846** (.000)
Attitude	.635** (.000)	1.000 .	.696** (.000)	.562** (.000)	.885** (.000)
Compliance	.614** (.000)	.696** (.000)	1.000 .	.301* (.013)	.796** (.000)
Open Mindedness	.535** (.000)	.562** (.000)	.301* (.013)	1.000 .	.742** (.000)

*Note.* \* =  $p \leq .05$ , \*\*  $p \leq .01$ .  $N = 54$  for all analyses. Significance appears in parentheses below correlations.

While the correlations amongst subscales fall into acceptable ranges, a factor analysis was conducted to determine how the questions on the TACO survey should be grouped. As shown in Appendix F, the factor analysis showed that there were 11 components with eigenvalues greater than 1. As depicted in Figure 2, the rotated component matrix shows that many of the derived subscales were similar to the intended subscales.

Rotated Component Matrix <sup>a</sup>				
	Component			
	1	2	3	4
T1		.692		
T2	.407	.534		
T3	.661			
T4	.808			
T5	.708			
T6	.758			
T7		.468		
T8			.494	
T9	.601		.500	
T10	.668			
A1	.574			
A2	.694			
A3				.783
A4			.434	.526
A5				.868
A6		.538	.486	.437
A7			.639	
A8			.599	
A9		.500	.531	
A10	.575		.489	

Rotated Component Matrix <sup>a</sup>				
	Component			
	1	2	3	4
C1	.589			
C2	.667			
C3			.555	
C4	.545		.449	
C5				.876
C6			.461	
C7			.831	.504
C8				
C9				
C10	.722			
O1		.534		
O2		.751		
O3		.779		
O4		.616		
O5		.473		
O6		.813		
O7		.773		
O8	.523			.408
O9		.590		
O10		.692		

Figure 2. Component matrix generated for factor analysis.

### Revised TACO Survey

Based on the results of the factor analysis, coupled with peer feedback, a revised version of the TACO test was developed with statements unique to each component. Furthermore, statements that loaded on multiple components, and were, therefore, interpreted differently amongst individuals, were removed from the TACO survey altogether. Items reflecting each subscale are no longer grouped together, and half of the items were reverse-coded to reduce response bias that occurred on the initial TACO. The revised TACO survey, shown in Table 4, consists of four demographic questions and 20 statements that require individuals to rate their behaviors and attitudes on a scale that ranges from *Strongly Agree* to *Strongly Disagree*; as shown in Appendix G, student responses are converted to numerical scores, where *Strongly Agree* = 5, *Neutral* = 3, and *Strongly Disagree* = 1. Each characteristic (Task Value, Attitude, Compliance, and Open-Mindedness) was measured with a series of five questions, two of which are reverse



scored, and subscale scores for each personality characteristic and a total TACO score are determined based upon student responses. It should be noted that the questions are not presented in subscales as was done in the original TACO survey; rather, questions are randomly sorted, as shown in Appendix H.

Table 4

*Revised TACO Survey*

Subscale	TACO Survey Questions
<b>Task Value</b>	<ol style="list-style-type: none"> <li>1) I think I will be able to use what I learn in one course in other courses.</li> <li>2) It is not important for me to learn the course material in a class.</li> <li>3) I think the course material in class is useful for me to learn.</li> <li>4) Understanding the subject matter of a course is not very important to me.</li> <li>5) I usually study in a place where I can concentrate on my course work.</li> </ol>
<b>Attitude</b>	<ol style="list-style-type: none"> <li>1) If I try hard enough, then I will understand the course material.</li> <li>2) It is my own fault if I don't learn the material in a course.</li> <li>3) If I don't understand the course material, it is because I didn't try hard enough.</li> <li>4) I'm certain I can understand the most difficult material presented in the readings for a course.</li> <li>5) It is not my fault if I don't learn material in a course.</li> </ol>
<b>Compliance</b>	<ol style="list-style-type: none"> <li>1) I memorize key words to remind me of important concepts in a class.</li> <li>2) I do not make lists of important terms for a course and memorize the lists.</li> <li>3) I make sure I keep up with the weekly readings and assignments for a course.</li> <li>4) When I become confused about something I'm reading for a class, I go back and try to figure it out.</li> <li>5) I do not keep up with the readings and assignments in a course.</li> </ol>
<b>Open Mindedness</b>	<ol style="list-style-type: none"> <li>1) When a theory, interpretation, or conclusion is presented in class or in the readings, I try to decide if there is good supporting evidence.</li> <li>2) I don't think about possible alternatives when I read or hear an assertion or conclusion in a class.</li> <li>3) In class, I prefer course material that doesn't really challenges me.</li> <li>4) I generally find that there are multiple ways to approach a given task.</li> <li>5) In class, I prefer course material that arouses my curiosity, even if it is difficult to learn.</li> </ol>

**Final Study**

For the final dissertation study, a convenience sample was used, which included 173

undergraduate education students at the University of Houston (UH) who are required to attend a Pedagogy and Professional Responsibilities (PPR) review session and complete any work associated with the review session. As shown in Figure 3, students were randomly assigned to one of three groups: one that participates in a content-based PPR review session, one that participates in a content-superseding technique (CST)-based PPR review session, and one whose scores reflect neither intervention, the control group. Pre-test scores were used, setting the control group's change at zero. While not optimal, ethical considerations precluded a control group with no review at all.

Participants completed the NEO-PI-R and TACO surveys a few weeks before the PPR intervention, and no relation between the two activities was suggested. The personality measures were introduced as providing individuals information that might help them understand themselves as teachers and were given as part of a required teacher training course. (Later in the semester, each student did receive an individual NEO profile report prepared by Psychological Assessment Resources, Inc., which owns the copyright.) The PPR intervention itself consisted of a 15-question PPR pretest (Appendix G), followed by a PPR review session, which used identical materials, but emphasized either content or techniques, immediately followed by an identical form 15-question PPR posttest. Unlike in the pilot study, the personality inventories, the pretest, and the posttest were administered via pencil and paper. The completed NEO-PI-R inventories were sent to Psychological Assessment Resources Inc., for scoring, while the TACO inventory, PPR pretest, and PPR posttest were hand-scored. All hand-scored items underwent cross-checks to ensure the accuracy of data entry.

Upon conclusion of the interventions, we analyzed the data to find correlations between both the overall and subscale TACO scores and NEO-PI-R scores and the score change between pre- and post-tests. Furthermore, we were able to determine the effect of intervention type on STS increases and whether students exposed to CSTs show more improvement in STS than those who are exposed to content-based interventions.

Study participants were informed of the voluntary nature of this study and that responses would be anonymized and not used against them. Furthermore, participants were informed of their rights as participants, and told that they would receive a NEO-PI-R personality analysis as a result of their participation. Informed consent from the participants was received. In order to ensure anonymity of our participants, participants used their University of Houston identification numbers, rather than names, on the various instruments. The research was approved by University of Houston's Institutional Review Board.

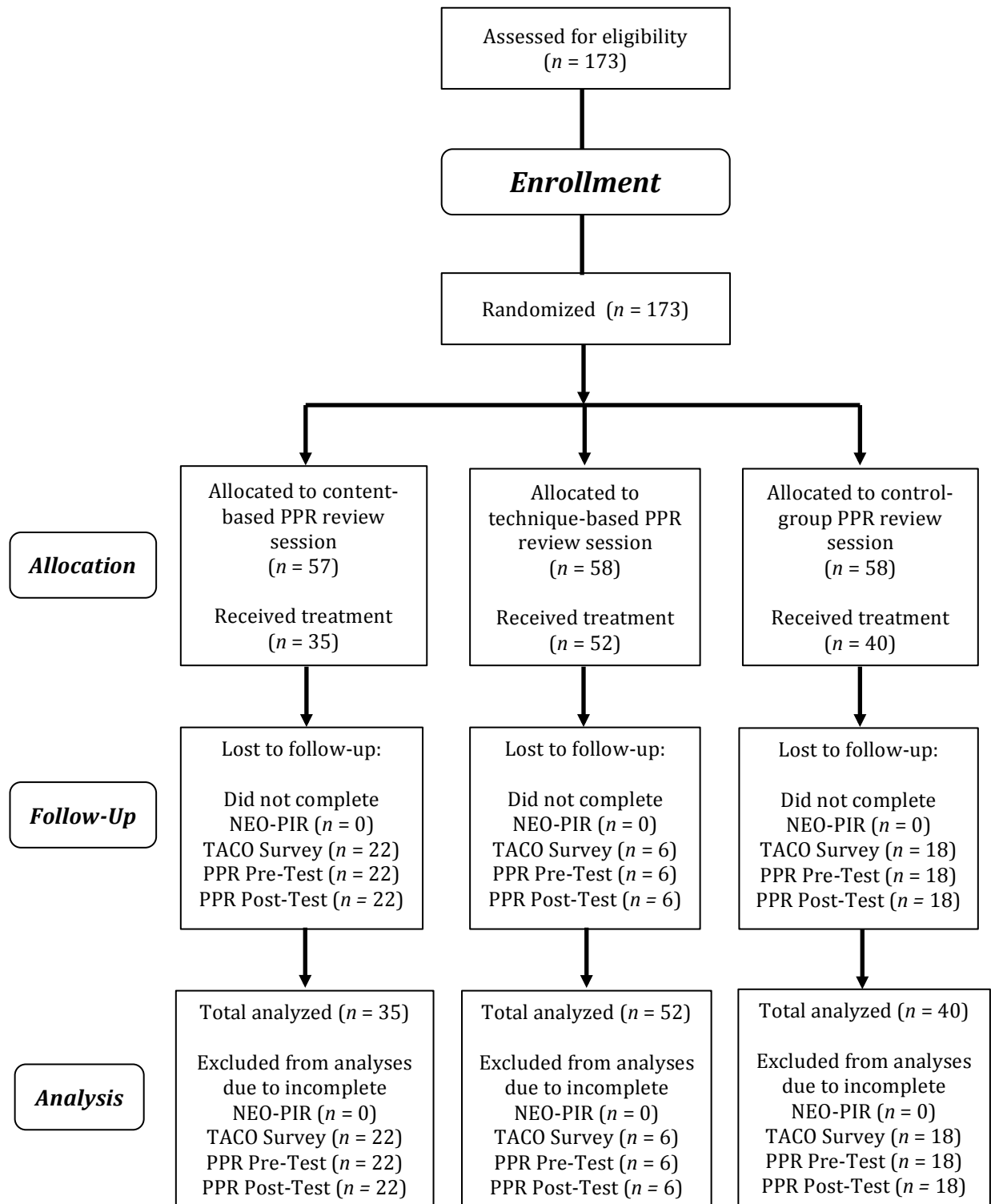


Figure 3. Flow of subjects diagram for final study

## **Chapter IV**

### **Results**

The aim of this study was to examine the relationships between the type of standardized-test (ST) intervention and standardized-test score (STS) increases and personality characteristics and STS increases. The study addressed the following research questions:

1. Do students achieve greater STS increases after a content- or technique-based intervention preparing for the PPR?
2. Do participants benefit differentially on PPR scores after content- or technique-based interventions?
3. Do the NEO-PI-R and TACO inventories assess similar or different characteristics?
4. Are personality variables on either the NEO-PI-R or the TACO survey associated with STS increases on the PPR?

Based on the information obtained from the above research questions, the ultimate question of the study is to clarify whether STs measure content mastery, mastery of CSTs, or personality variables.

### **Data Analysis**

This was a quantitative, experimental study that analyzed data from the NEO-PI-R, TACO inventory, PPR pretest, and PPR posttest. The completed NEO-PI-R inventories were sent to Psychological Assessment Resources Inc., for scoring, while the TACO inventory, PPR pretest, and PPR posttest were hand-scored. After collecting all

the data, it was entered into the Statistical Package for Social Sciences (SPSS) 22.0 for analysis.

In order to determine whether relationships existed among ST preparation intervention, STS increase, and personality characteristics a number of statistical tests were conducted. Descriptive statistics were generated for the demographic questions, NEO-PI-R domains, TACO scales, and score change between the PPR pre- and posttests, and included the means and standard deviations for each of the 127 responses. The means for score change were then analyzed using a univariate ANOVA, which measures the influence of an independent variable on a dependent variable, to determine the significance of score change by intervention. The means for personality factors on both the NEO-PI-R and TACO surveys were analyzed using a Pearson correlation matrix to determine whether overlap exists between personality factors measured on both personality surveys. The means for personality factors on both the NEO-PI-R and TACO surveys and score change were analyzed using a correlation matrix to determine the significance that personality factors have on ST score improvement. Given the exploratory nature of the research questions, statistically significant relationships were determined based on an alpha level of 0.10 or less (Stevens, 1986).

### **Demographic Data**

Both the NEO-PI-R and TACO surveys contained questions used to gather demographic data about the participants in the study. A total of 127 participants took part in the study, of which 103 (82.4%) were women and 22 (17.6%) were men.

The age of the 127 participants ranged from 19 to 44 years ( $M = 22.55$ ,  $SD = 4.21$ ). Age was non-normally distributed with a skewness of 2.95 ( $SE = 0.22$ ) and

kurtosis of 10.24 ( $SE = 0.43$ ). Participants were also asked to provide their intended teacher certification level. 69 (55.2%) participants were pursuing their elementary certification, 34 (27.2%) participants were pursuing their middle school certification, and 22 (17.6%) participants were pursuing their high school certification.

### NEO-PI-R, TACO, and PPR Intervention Results

Descriptive statistics of mean and standard deviation were calculated for score change by intervention, as shown in Table 5, and for each subscale of both the NEO-PI-R and TACO inventories, as shown in Table 6.

Table 5

#### *Score Change by Intervention*

	<i>N</i>	<i>M<sub>change</sub></i>	<i>SD</i>
Intervention Type			
Control	40	0.00	0.00
Content	35	0.11	1.97
Content-Superseding Technique	52	0.67	1.61

Table 6

#### *TACO and NEO-PI-R Subscale Scores*

Measure	<i>M</i>	<i>SD</i>
TACO		
Task Value	4.46	0.43
Attitude	3.73	0.54
Compliance	4.13	0.49

---

Open-Mindedness	4.05	0.45
NEO-PI-R		
Neuroticism	55.28	11.05
Extraversion	54.04	10.51
Openness to Experience	56.88	10.55
Agreeableness	54.58	10.15
Conscientiousness	55.50	11.52

---

*Note.*  $N = 127$  for all analyses.

### **Score Change by Intervention**

A one-way between-groups analysis of variance was conducted to explore the effect of intervention type on score change, as measured by the dependent variable score change, which was calculated as the number of questions a participant answered correctly on the Pedagogy and Professional Responsibilities (PPR) posttest subtracted from the number of questions a participant answered correctly on the PPR pretest. Participants had been randomly divided into three intervention groups (Group 1: Control; Group 2: Content-Based; Group 3: Technique-Based). There was a statistically significant difference at the  $p < 0.10$  level in score change for the three intervention groups:  $F(2, 124) = 2.84, p = 0.06$ . Furthermore, the actual difference in mean score change between the content-based intervention and technique-based intervention was relatively large. The effect size of the gain, calculated using Cohen's  $d$ , was 0.34, a medium effect size. Post-hoc comparisons using the Tukey HSD test indicated that the mean score change for Group 1 ( $M = 0, SD = 0$ ) was significantly different from Group 3 ( $M = 0.67, SD = 1.61$ ),



and the mean score change for Group 2 ( $M = 0.11$ ,  $SD = 1.96$ ) was significantly different from Group 3, as shown in Figure 5.

### Score Change by Personality Characteristics

To determine the relationship between score change and personality characteristics, both the TACO and NEO-PI-R surveys were administered to the 127 participants.

First, it was necessary to analyze the relationship between the subscales of the TACO Survey. The relationship between the personality variables of Task Value, Attitude, Compliance, and Open-Mindedness, as measured by the TACO Survey, was investigated using Pearson product-moment correlation coefficient. The results of this correlation are shown in Table 7. It should be noted that the high correlation among scales on the TACO inventory indicates significant subscale overlap, and further development is required to consider the TACO inventory a valid and reliable scale.

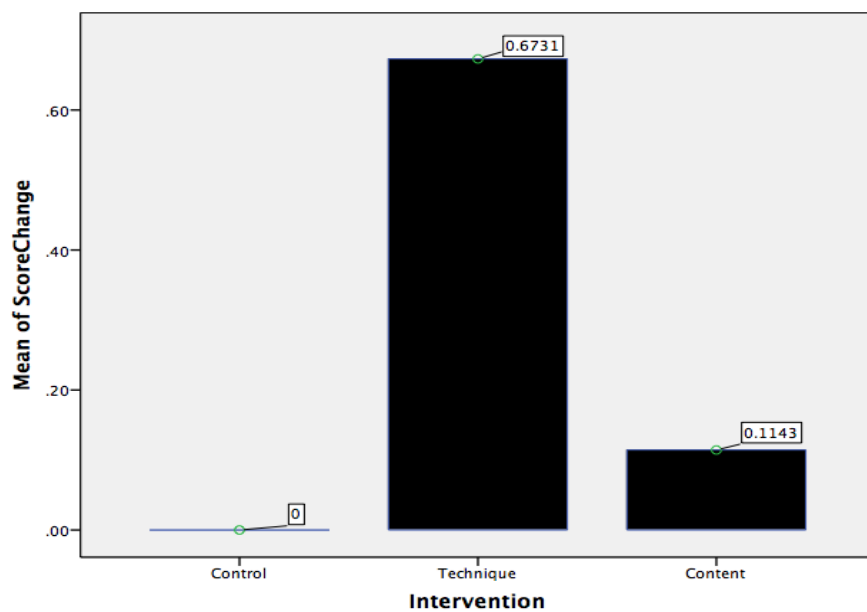


Figure 4. Mean difference in score change by intervention type

Table 7

*Pearson's Correlations Between Personality Characteristics on the TACO Survey*

	Personality Characteristics on TACO			
	Task Value	Attitude	Compliance	Open-Mindedness
Task Value	1.000 .			
Attitude	.307** .	1.000 .		
Compliance	.438** (.000)	.149 (0.95)	1.000 .	
Open Mindedness	.415** (.000)	.280** (.001)	.472** (.000)	1.000 .

*Note.* \*\* =  $p \leq .01$ .  $N = 127$  for all analyses. Significance appears in parentheses below correlations.

Next, the relationship between the personality variables of Task Value, Attitude, Compliance, and Open-Mindedness, as measured by the TACO Survey, and the personality variables of Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness, as measured by the NEO-PI-R, was investigated using Pearson product-moment correlation coefficient. The results of this correlation are shown in Table 8. Of particular note is the correlation between the Conscientiousness scale of the NEO-PI-R and all the TACO scales; i.e., the TACO scales of Task Value, Attitude, Compliance, and Open-Mindedness are all strongly related to the Conscientiousness scale on the NEO-PI-R.

Table 8

*Pearson's Correlations Between TACO and NEO-PI-R Personality Subscales*

	Personality Characteristics on TACO			
	Task Value	Attitude	Compliance	Open-Mindedness
Neuroticism	-.035 (.699)	-.170 (.056)	-.177** (.047)	-.286** (.001)
Extraversion	.087 (.332)	.194* (.029)	.100 (.264)	.172 (.054)
Openness to Experience	.134 (.134)	.071 (0.429)	.146 (.103)	.431** (.000)
Agreeableness	.099 (.269)	.099 (.268)	.085 (.343)	.107 (.231)
Conscientiousness	.428** (.000)	.275** (.002)	.469** (.000)	.260** (.003)

*Note.* \* =  $p \leq .05$ , \*\*  $p \leq .01$ .  $N = 127$  for all analyses. Significance appears in parentheses below correlations.

Finally, the relationship between STS change and personality variables measured by both the NEO-PI-R and TACO inventories was analyzed using Pearson product-moment correlation coefficient. The results of this correlation are shown in Table 9.

Table 9

*Pearson's Correlation Between Score Change and TACO and NEO-PI-R Domains*

Personality Variable	Score Change
NEO-PI-R	
Neuroticism	.120 (.179)
Extraversion	.007 (.939)
Openness to Experience	-.202* (.023)
Agreeableness	.067 (.452)
Conscientiousness	-.171 (.054)
TACO	
Task Value	-0.50 (.579)
Attitude	-.138 (.122)
Compliance	-.142 (.112)
Open-Mindedness	-.078 (.382)

*Note.* \* =  $p \leq .08$ .  $N = 127$  for all analyses. Significance appears in parentheses below correlations.

## **Chapter V**

### **Discussion and Conclusion**

The aim of this study was to examine the relationships between the type of standardized-test (ST) intervention and standardized-test score (STS) increases and personality characteristics and STS increases. The ability to increase STSs with tactical test preparation indicates that ST questions can be answered with content-superseding techniques (CSTs) rather than content knowledge, thus negating the validity of a ST as an objective measure of student aptitude or achievement. Since test-preparation companies are unwilling to provide information regarding their trade secrets, there is a dearth of research regarding the benefits of using CSTs instead of content knowledge on STs. Furthermore, there is a lack of research on the personality traits present in individuals who attain significant STS increases using CSTs. However, previous studies have indicated that an individual's personality traits have a significant impact on academic achievement, subject-specific test scores, and the way in which learning occurs. In order to determine whether or not an individual can achieve greater STS increases using CSTs rather than content knowledge, and the personality traits present in individuals who do succeed using CSTs, the study addressed the following research questions:

1. Do students achieve greater STS increases after a content- or technique-based intervention preparing for the PPR?
2. Do participants benefit differentially on PPR scores after content- or technique-based interventions?
3. Do the NEO-PI-R and TACO inventories assess similar or different characteristics?

4. Are personality variables on either the NEO-PI-R or the TACO survey associated with STS increases on the PPR?

Using the information garnered from the aforementioned research questions, the ultimate question of the study is to clarify whether STs measure content mastery, mastery of CSTs, or personality variables.

The following discussion will address the aforementioned research questions, presenting the data acquired through the intervention, the implications of the findings, the strengths and weaknesses of the study, and future directions.

### **Standardized Test Score Increases by Intervention Type**

In order to address the first two research questions, and determine whether individuals achieve greater standardized test score increases after exposure to content-based or content-superseding technique (CST) based learning interventions and whether individuals benefit differentially from content-based or CST-based learning interventions, a three-group, randomly assigned, pretest-posttest experiment was conducted. As stated in the methodology, 173 participants were randomly assigned to three groups for the PPR Review Sessions: a control group, a content-based intervention group, and a CST-based intervention group. Due to loss of subjects between administration of the NEO-PI-R and TACO personality inventories and PPR interventions, data from a total of 127 participants were evaluated using a one-way between-groups analysis of variance to assess the effect of intervention type on score change.

Analysis of the data found that the difference in score change at the  $p < 0.10$  level between the content-based intervention and CST-based intervention was relatively large, resulting in a medium effect size of 0.34, calculated using Cohen's  $d$ . The data indicated

that the mean score change for the control group ( $M = 0$ ,  $SD = 0$ ) was significantly different from the CST-based group ( $M = 0.67$ ,  $SD = 1.61$ ), and the mean score change content-based group ( $M = 0.11$ ,  $SD = 1.96$ ) was significantly different from CST-based group. Specifically, the average STS increase was over 600% greater for the CST-based group than for content-based group. Either intervention fared better than no intervention. Since all PPR interventions consisted of 90 minutes of instruction, and test-preparation companies offer preparation courses ranging from 6-30 hours in length, the implicit potential to further increase STS using CSTs is massive. Thus, the study revealed a significant relationship between STS increases and CST-based interventions.

Furthermore, the data confirms previous research that STs do not assess an individual's mastery of content knowledge. According to ETS (2017), the content and reasoning abilities assessed in STs reflect the critical thinking skills necessary for success in a particular arena; in the case of the TExES PPR Exam, the abilities assessed supposedly reflect critical thinking skills necessary for successful teaching in the state of Texas. Thus, many institutions believe that the PPR accurately predicts whether or not an individual is suited for, and will be successful in, a K-12 classroom (Kuncel & Hezlett, 2007). Unfortunately, the use of the TExES PPR Exam, or any ST, as a metric for gauging student knowledge and ability is problematic due to test-takers' ability to neglect the content by using CSTs. Thus, the use of STs as a predictor of content knowledge is flawed, as individuals can achieve STS increases without content mastery (Koretz, 2013; Popham, 1999).

Due to the paucity of research comparing the effect of content-based and CST-based interventions, the results of this study cannot be related to previous research.

However, the findings do support the concept that test-preparation companies can prove beneficial to individuals who are willing and able to participate in test-preparation courses. The ever-burgeoning test-preparation industry has capitalized on the learnability of STs by offering a plethora of courses aimed at overcoming the difficulties many individuals face when taking STs. Indeed, the continued growth of the test-preparation industry is supported by the findings of this study that show that CST-based interventions can lead to significant STS increases, allowing test-preparation companies to offer money-back guarantees should individuals not achieve desired score increases (Kaplan, 2017; The Princeton Review, 2017b; Testmasters, 2017).

### **Standardized Test Score Increases and Personality Variables**

To determine whether personality variables on either the NEO-PI-R or the TACO inventory associated with STS increases,  $N = 173$  participants took both the NEO-PI-R and TACO personality inventories. Due to loss of subjects between the administration of the personality inventories and PPR interventions, data from 127 participants were analyzed.

The first stage of this analysis involved a Pearson correlation between subscales on the NEO-PI-R inventory, Neuroticism, Extraversion, Openness to Experience, Agreeableness, and Conscientiousness, and the subscales on the TACO inventory, Task Value, Attitude, Compliance, and Open-Mindedness, to determine a relationship between scales on the personality inventories. Perhaps the most significant finding was that all of the subscales on the TACO inventory were significantly correlated with the subscale of Conscientiousness on the NEO-PI-R. Such a strong correlation indicated that the TACO inventory was not, in fact, analyzing four different scales, but instead was assessing



different factors of Conscientiousness. In order to determine what factors of Conscientiousness are actually assessed on the TACO inventory, it is imperative that future researchers conduct a factor analysis between facets of the NEO-PI-R's Conscientiousness domain and items on the TACO inventory. Based on the conducted analysis, as written, the TACO inventory is not a valid scale of an individual's belief in the Task Value of, Attitude towards, Compliance with, or Open-Mindedness to an assignment or course of study.

The second phase of this analysis involved conducting a Pearson correlation between scales on the NEO-PI-R and TACO personality inventories and score change attained from the PPR Review intervention. The only significant correlation occurred between score change and the NEO-PI-R scale of Openness to Experience, where  $r = -.202$ , indicating that a score increase was negatively correlated with an individual's openness to experience. While an initial view may dismiss this finding, one must consider what the domain of Openness to Experience measures. As defined by Costa and McCrea (1992), the domain of Openness to Experience assesses an individual's active imagination, aesthetic sensitivity, attentiveness to inner feelings, preference for variety, intellectual curiosity, and independence of judgment. Furthermore, Openness to Experience is especially related to aspects of intelligence, such as divergent thinking (McCrae, 1987), and those who score low on this domain tend to be conventional in behavior.

Considering that ST questions can be overcome by using CSTs, the negative correlation between Openness to Experience and STS increase supports the idea that individuals who are less willing to adopt alternative test-taking strategies are less likely to improve STS. In essence, an individual who is less likely to think critically, instead

favoring techniques that allow one to undermine the content knowledge assessed, is more likely to obtain a significant score increase on a standardized test.

### **What Do Standardized Tests Measure?**

The ultimate quandary explored in this research was whether STs measure content mastery, content-superseding technique mastery, personality variables, or a combination of the aforementioned. Based on the analyses conducted, it is evident that the content knowledge that STs attempt to measure can be made irrelevant by limiting critical thinking skills (reflected by the Openness to Experience scale on the NEO-PI-R), and using content-superseding techniques (CSTs), which are based on word-identification rules for choosing right answers. This finding is supported by professor William Ayers (2001), who notes that, “standardized tests can't measure initiative, creativity, imagination, conceptual thinking, curiosity, effort, irony, judgment, commitment, nuance, good will, ethical reflection, or a host of other valuable dispositions and attributes” (p. 112). Indeed, individuals who eschew the characteristics described by Ayers were more likely to achieve significant STS increases than those who embrace such complex traits.

### **Implications of Study Results**

The implications of tactical test preparation as a means of improving standardized-test scores (STS) are troublesome, at best. Such significant score improvement after a 90-minute intervention indicates that the TExES Pedagogy and Professional Responsibilities (PPR) exam assesses learnable, lower-level reasoning skills (such as identifying rule-out words and testing formulas by trial and error substitution) and, therefore, is an inadequate predictor of success in the field of education. It should be noted that the PPR is developed and administered by Educational Testing Services (ETS),

the same company that is responsible for creating and administering college admissions and placement exams such as the Preliminary SAT (PSAT), SAT, Test of English as a Foreign Language (TOEFL), College Level Exam Placement (CLEP), and Advanced Placement assessments. Furthermore, ETS is responsible for graduate school admissions exams such as the Graduate Record Examination (GRE) and the Praxis test, as well as state-mandated exams such as the State of Texas Assessments of Academic Readiness (STAAR), the California Standardized Testing and Reporting (STAR) program, and the National Assessment of Educational Progress (NAEP).

The ability to undermine the content tested on these exams with CSTs supports the idea espoused by Brady (2011) that standardized tests (STs) cover a limited amount of subject matter in a superficial manner and fail to address higher-level critical thinking skills. Unlike many tests encountered in educational settings, STs are composed by skilled psychometricians, tested, refined, and retested in an attempt to ensure questions are both fair and objective (Phelps, 2002). The aforementioned process of test development creates a predictable form and structure of ST questions—a form and structure that can be learned and exploited using CSTs. Accordingly, individuals can obtain significant STS increases without a thorough understanding of the content the test intends to assess. Thus, despite attempts to create a valid measure for comparison amongst peers, STs fail to provide meaningful insight about an individual's knowledge (Popham, 1999); the PPR is no exception.

One must also consider the implications of using STS as a metric of student, educator, and institutional achievement. In 1969, the Department of Education (DOE) created the National Assessment of Educational Progress test to identify the educational

accomplishments of students in the public institutions (Kumeh, 2011), and the Elementary and Secondary Education Act (ESEA) of 1965 authorized the use of federal funds to support K-12 education programs (Standerfer, 2006). The continued emphasis of legislation, including No Child Left Behind (NCLB) and Race to the Top (RTTT), has led to an increased emphasis on testing in school, and a decline in curriculum-based education.

While RTTT was formed with the intention of improving the nature of education, the legislation has led to further problems. In order to receive funding, schools must meet certain criteria, much of which is measured by standardized test scores. Under RTTT, not only do ST scores claim to measure student success, but also they are directly linked to teacher performance evaluations. Accordingly, educators and administrators are placing increasing importance on ST in the classroom since their employment often depends on student performance (Dexheimer, 2009). Indeed, in his announcement of RTTT, Secretary Arne Duncan (2009) stated, “Once new standards are set and adopted, you need to create new tests that measure whether students are meeting those standards. How can you possibly talk about teacher quality without factoring in student achievement?”

In turn, since the enactment of NCLB and RTTT, and to the detriment of all involved, ST preparation has become saturated into all aspects of public school curricula. The increased emphasis on test preparation in the classroom has led to a decrease in student knowledge and places an undue burden on educators responsible for student STS; in order to ensure continued employment, educators place increased focus on STs, sacrificing curricula. Rather than providing students with an enriching educational experience, educators have begun to focus their efforts on teaching for STs. Indeed, the

director of the Center for Education at Rice University. Linda McNeil (2002), observes that the classroom now consists of, “phony curricula, reluctantly presented by teachers in class to conform to the forms of knowledge their students would encounter on centralized tests” (p. 5).

The current emphasis on STs detrimentally affects all actors involved in the education system, including students, teachers, and administrators. Educators are “teaching to the test,” excessively drilling students over ST material, and promoting rote memorization as a major form of learning (Crocco & Costigan, 2007; Jennings & Bearak, 2014; McNeil, 2002; Nelson, 2013). According to a 2007 study by the Center of Education policy, 44% of school districts have eliminated two and a half hours of science, arts, and social studies instruction each week in order to focus on math and reading that will occur on STs (Ravitch, 2010). If all is to remain status quo, the education system will no longer represent a place of learning and knowledge wherein innovation and critical thought is encouraged. Rather, schools will simply become places where one can learn how to take a ST. Furthermore, based on the study results, educators could increase student STS by emphasizing CSTs rather than content knowledge, securing their employment, but promoting lower-level critical thinking skills.

Not only does the ability to achieve greater STS increases with CSTs undermine the purpose of both STs and education at large, but also the cost of test-preparation courses precludes those of a lower socioeconomic status from participating in such courses. Test-preparation courses are quite expensive and, therefore, are limited to individuals in the upper echelons of society; indeed, underserved populations are at a stark disadvantage when taking STs due to the cost and accessibility of test-preparation

courses. As of Spring 2017, a course at a test-preparation company can vary from 7699-\$10,000, depending on the company used, test selected, and number of hours of preparation desired (The Princeton Review, 2017; Kaplan, 2017; Testmasters, 2017). Despite the fact that many students may wish to partake in a test-preparation course, the financial burden is often prohibitive. While the affluent reap the rewards associated with such comprehensive test-preparation, underserved populations, or individuals unable to handle the financial burden associated with ST preparation courses, are at a stark disadvantage when compared with those who can afford such directed instruction. While test-preparation companies release texts that aim to instruct individuals on CSTs, individuals are limited to the scope of such texts and do not benefit from the direct interaction that occurs in test-preparation courses. Furthermore, the burden of understanding the purpose and methods of CSTs lies with individuals who purchase such texts; i.e., students must be motivated to learn the CSTs on their own without the help of an instructor who is a subject-matter expert.

Indeed, the implications of the successful use of CSTs to raise STS is concerning when one considers the accessibility of test-preparation services: the apparent objective, unbiased nature of a ST is degraded, and underserved populations are unknowingly, and unfairly, compared to individuals who can afford test-preparation services. Therefore, to consider the PPR, or any ST, a predictor of student success or measure of intelligence, in any arena, is both flawed and inaccurate (Baker, Bartion, Darling-Hammond, Haertel, Ladd, & Linn, 2012). Therefore, the current use of STs must be altered in one of two ways: tests must be developed that overcome the CSTs professed by test-preparation

companies or STs must be eliminated altogether to allow equal opportunities for all individuals.

However, since institutions and professional associations use STs as a quantitative metric for student comparison, it is improbable that STs will be eliminated altogether. In order to overcome the problematic nature of STs, but still provide a metric for student comparison, institutions could create and administer exams that specifically address the specialties and knowledge valued by that particular program or institution. Rather than requiring the TExES PPR as a necessary component of the teacher certification process, educational institutions should consider creating alternative performance assessments (Koretz, 2013; Neal, 2013), unique to each school, that gauge potential educators against a pre-determined set of criteria. In essence, each institution could create its own criterion-referenced ST, to determine whether an applicant is a suitable fit for a particular institution. Likewise, institutions of higher education could develop and administer admissions exams that reflect the values and requirements of a specific institution, rather than depending on STs such as the SAT or GRE. After all, the idea that a single ST can be used to determine whether an individual is a solid fit for institutions worldwide is problematic; a one-size-fits-all approach neither addresses the particular needs of an institution nor does it emphasize the strengths and limitations of the individual.

### **Strengths and Limitations**

As with all research studies, this study has a number of strengths and limitations. A major strength of this research is that the experimental design of the study enables one to draw a causal conclusion regarding ST interventions. Threats to internal validity were minimized in a number of ways. Primarily, all participants were in the Developing

Teaching phase of the Teacher Education Program at the University of Houston (UH), and were randomly assigned to the control, content-based, or CST-based intervention groups in order to control for subject characteristics and selection threats. To minimize history, maturation, and location threats, all participants completed the NEO-PI-R and TACO inventories during new student orientation and the PPR Review over a four-day period in similar classrooms in the College of Education building at the UH.

Additionally, the PPR pre- and posttest consisted of questions previously released by ETS, ensuring the validity and reliability of the exam questions; similarly, the NEO-PI-R is a verified psychological inventory. As noted above, experimenter effects were minimized by having one expert in both deliveries perform both reviews. While the study did lose  $N = 46$  participants before completion, the sample size was still large enough, at  $N = 127$ , to make significant findings and generalize the results to similar populations.

Testing threats were also minimized by the nature of the intervention itself; all participants completed identical forms of the NEO-PI-R and TACO inventories and identical form pre- and posttests were used in all PPR intervention groups, without allowing students to know their pretest scores or review the questions and answers in between the pre- and posttest. In order to reduce implementation threats, one individual was responsible for administering the NEO-PI-R and TACO inventories, while a separate instructor, versed in both content knowledge and CSTs applicable to the PPR, delivered all the PPR review sessions, and the personality and PPR tests appeared separate in timing and purpose to the students. Finally, it should be noted that all data obtained from participants was anonymized, so that participants would not tie participation in the intervention to beneficial or detrimental treatment in the UH Teacher Education Program.



However, while the results of this study are significant, it should be noted that there are a few limitations to this study. One limitation of this study involves the research design itself. While the before-and-after design may offer more valid results than other, non-experimental designs, the potential impact of pre-testing can limit the generalizability of the study if participants become aware of the experimental variable; future studies should account for this potential effect by using the Solomon Four-Group Design. Furthermore, the instructor who led the control, content-based, and CST-based interventions had experience working in both the traditional classroom setting and the test-preparation industry. Accordingly, it is feasible that the instructor displayed preference for one of the intervention methods, leading to the study results. Finally, despite informing participants that participation in this study would be anonymized and, in turn, have neither a beneficial nor detrimental effect on their program standing, it is possible that scores on both the NEO-PI-R and TACO personality inventories are artificially affected due to the Hawthorne Effect.

### **Future Directions**

While our study showed that students significantly improved STS on the TExES Pedagogy and Professional Responsibilities (PPR) exam using knowledge of CSTs rather than knowledge of content, considerable work remains in this arena. Perhaps most importantly, it is necessary to verify the results gleaned in this study by replicating the results in a future study. Future researchers should consider conducting a similar TExES PPR study at multiple higher education institutions in Texas that have similar teacher education programs to the one at the UH. Should researchers find similar results, the study findings could be generalized to a larger population of individuals.

Similarly, future research should conduct interventions and analyze data for other standardized examinations. Indeed, since the use of CSTs showed improvement on the PPR exam, an exam that caters to a limited population and for which test-preparation courses are not widely offered, it is of particular import to determine scientifically whether CSTs can be successfully employed on more common, widely-taken exams, such as the SAT or GRE. While test-preparation companies tout satisfaction guarantees if score improvements are not achieved, the industry is notoriously secret regarding its processes and does not release any data regarding score improvements. Furthermore, these industries do not allow their employees to analyze company-obtained student data or use such data for research purposes. Therefore, future research should involve conducting similar research that determines the power of CSTs on a previously released SAT or ACT exam.

Furthermore, future research should consider lengthening the time of each intervention. The varied interventions, which consisted of a pretest, PPR Review, and posttest, lasted a total of two hours each. Accordingly, only cursory information could be covered in both the content-based and CST-based interventions regarding the PPR exam. It would be particularly interesting to see how score change is affected by a lengthier intervention. Thus, future research should conduct longer review periods, or multiple review periods, between the pretest and posttest.

Finally, in order to glean more information regarding the relationship between personality characteristics and STS increases, future researchers should consider analyzing the facets of Conscientiousness on the NEO-PI-R against score change to see if particular facets of Conscientiousness specifically relate to STS increases. If the facets of

the Conscientiousness factor align with different subscales on the TACO inventory, the NEO-PI-R could be used to determine whether or not a student will benefit from using CSTs on standardized tests. Conversely, if the subscales of the Conscientiousness factor do not align with different subscales on the TACO inventory future research would involve revision of the TACO inventory.

## **Conclusion**

While the debate over the validity and reliability of STs will, undoubtedly, continue for some time, it is clear from the data that STs assess learnable skills. As the data from the PPR Review sessions show, participants were able to achieve significantly higher score improvement by shunning critical thinking skills and complying with skills espoused in an intervention focused on improving STS with CSTs. Such findings explain the growth and sustainability of test-preparation services that afford individuals with the opportunity to improve STS by emphasizing the use of CSTs that undermine content knowledge.

Damage reversal and effective change, however, is possible. Due to the issues associated with STs such as the PPR, it is imperative that an alternative means of assessment be used in all arenas that rely on such exams. With regards to the PPR exam, the use of CSTs allows future educators to obtain a teaching certificate without having a firm grasp on the pedagogical concepts the PPR intends to assess. One way to improve the current situation would involve the cessation of standardized testing in favor of more objective performance assessments. Arguably, institutions can glean far more valuable quantitative and qualitative information from transcripts, personal statements, résumés, and recommendations, and, therefore, should not base an individual's qualifications on

STS. After all, the rich nature of the qualitative data, and the mass of quantitative data provided by transcripts, provide insight to an individual's strengths and weaknesses—something that a ST cannot provide.

Institutions and organizations that utilize STS as a determining factor in admissions and employment processes are, essentially, rewarding individuals for effective use of lower-level thinking skills. Eliminating STs, or developing alternative performance assessments, would rectify four issues: students would be required to utilize higher-level thinking skills to demonstrate content mastery, students would be assessed on material that is in line with the institutional curriculum, students would not be penalized should they come from an underserved population, and educators would be held accountable for teaching content rather than lower-level thinking skills.

## References

- Akers, R. P. (2016). *The utility of a standards-based teacher evaluation as a measure of effectiveness* (Doctoral dissertation, Frostburg State University).
- Aliaga, M. and Gunderson, B. (2005). *Interactive statistics*. Upper Saddle River, New Jersey: Prentice.
- Ayers, W. (2001). *To teach: The journey of a teacher*. New York, NY: Teachers College Press.
- Baker, E. L., Bartion, P. E., Darling-Hammond L., Haertel, E., Ladd, H. F., Linn, R. L., et al. (2010, August). *Problems with the use of student test scores to evaluate teachers: EPI briefing paper # 278*. Economic Policy Institute.
- Bond, Linda A. (1996). Norm- and criterion-referenced testing. *Practical Assessment, Research & Evaluation*, 5, Retrieved from <http://PAREonline.net/getvn.asp?v=5&n=2>.
- Brady, M. (2011, November 1). The complete list of problems with high-stakes standardized tests. *The Washington Post*. Retrieved from [http://www.washingtonpost.com/blogs/answer-sheet/post/the-complete-list-of-problems-with-high-stakes-standardized-tests/2011/10/31/gIQA7fNyaM\\_blog.html](http://www.washingtonpost.com/blogs/answer-sheet/post/the-complete-list-of-problems-with-high-stakes-standardized-tests/2011/10/31/gIQA7fNyaM_blog.html)
- Brody, B. (2009, October 29). Does SAT prep really help? *Westchester Magazine*. Retrieved from <http://www.westchestermagazine.com/Westchester-Magazine/November-2009/To-Prep-or-Not-to-Prep/>
- Center on Education Policy. (2008). *Has student achievement increased since 2002? State test score trends through 2006-07*. Washington, DC: Author.

- Center on Education Policy. (2009). *Has student achievement increased since 2002? State test score trends through 2006-07*. Washington, DC: Author.
- Chamorro-Premuzic, T. & Furnham, A. (2003). Personality traits and academic examination performance. *European Journal of Personality*, 17, 237-250.
- Coffey, A., & Atkinson, P. (1996). Narrative and stories. In *Making sense of qualitative data: Complementary research strategies* (pp. 54-82). Thousand Oaks, CA: Sage.
- Conard, M. A. (2006). Aptitude is not enough: How personality and behavior predict academic performance. *Journal of Research in Personality*, 40, 339-346.
- Costa, P.T., Jr. & McCrae, R.R. (1992). *Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI) manual*. Odessa, FL: Psychological Assessment Resources.
- Crocco, M. S. & Costigan, A. T. (2007). The narrowing of curriculum and pedagogy in the age of accountability: Urban educators speak out. *Urban Education*, 6, 512-535.
- Davis, C. E. (2008). *Regression to the mean*. Wiley Encyclopedia of Clinical Trials. 1–2.
- Dexheimer, E. (2009, March 19). The high cost of TAKS. Retrieved from [http://www.statesman.com/blogs/content/shared-gen/blogs/austin/investigative/entries/2009/03/19/the\\_high\\_cost\\_of\\_taks.html](http://www.statesman.com/blogs/content/shared-gen/blogs/austin/investigative/entries/2009/03/19/the_high_cost_of_taks.html)
- DiCicco-Bloom, B. & Crabtree, B. F. (2006). The qualitative research interview. *Medical Education*, 40, 314-321.
- Duncan, A. (2009, June 14). States will lead the way toward reform. Keynote address presented at the 2009 Governors Education Symposium, Cary, NC. Retrieved from: <http://www.ed.gov/news/speeches/states-will-lead-way-toward-reform>

- Educational Testing Services. (2017). About the e-rater scoring engine. Retrieved from <http://www.ets.org/erater/about>
- FairTest-National Center for Fair & Open Testing. (2017). *Norm-referenced achievement tests*. Retrieved from website <http://www.fairtest.org/sites/default/files/norm-referenced-tests.pdf>
- Fisher, R. A. (1973). *Statistical methods and scientific inference*. London: Collins Macmillan.
- Garland, S. (2012, October 9). Online standardized tests are new, but are they revolutionary? *The Huffington Post*. Retrieved from [http://www.huffingtonpost.com/2012/10/10/online-standardized-tests\\_n\\_1954569.html](http://www.huffingtonpost.com/2012/10/10/online-standardized-tests_n_1954569.html)
- Hickok, E., & Ladner, M. (2007, June 27). *Reauthorization of No Child Left Behind: Federal management or citizen ownership of K-12 education?* (Research Report 2047). Retrieved from <http://www.heritage.org/Research/Reports/2007/06/Reauthorization-of-No-Child-Left-Behind-Federal-Management-or-Citizen-Ownership-of-K-12-Education>
- Hofer, M., Kuhnle, C., Kilian, B., & Fries, S. (2012). Cognitive ability and personality variables as predictors of school grades and test scores in adolescents. *Learning and Instruction, 22*, 368-375.
- Hout, M., Elliott, S., & Freuh, S. (2012). Do high-stakes tests improve learning? *Issues in Science & Technology, 29*, 33-38.
- Hosp, J. L., Drew, C. J., & Hardman, M. L. (2007). *Designing and conducting research in education*. Thousand Oaks, CA: Sage.

- Howard, C. (2014, March 5). Right answer for SAT? College board revamps test, partners with Khan Academy. Retrieved from <http://www.forbes.com/sites/carolinehoward/2014/03/05/right-answer-for-sat-college-board-revamps-test-adds-partnership-with-khan-academy/>
- Hull, J. (2013). Trends in teacher evaluation: How states are measuring teacher performance. Retrieved from <http://www.centerforpubliceducation.org/Main-Menu/Evaluatingperformance/Trends-in-Teacher-Evaluation-At-A-Glance/Trends-in-Teacher-EvaluationFull-Report-PDF.pdf>
- Jennings, J. L. & Bearak, J. M. (2014). "Teaching to the test" in the NCLB Era: How test predictability affects our understanding of student performance. *Educational Researcher*, 43, 381-389.
- Kaplan Test Preparation. (2017). *Higher score guarantee*. Retrieved from <http://www.kaptest.com/hsg>
- Kaplan Test Preparation. (2017b). *SAT Prep Course*. Retrieved from <https://www.kaptest.com/sat/sat-prep-course/sat-classroom>
- Kastenbaum, S. (2012, May 17). The high stakes of standardized tests. *CNN*. Retrieved from <http://schoolsofthought.blogs.cnn.com/2012/05/17/the-high-stakes-of-standardized-tests/>
- Kiesler, C. A. (1998, Feb 25). Affirmative action and the SAT. *Education Week*, 17, 60.
- King, F. J., Goodson, L., & Rohani, F. (1998) Higher-order thinking skills: Definitions, teaching strategies, and assessment. Retrieved from [http://www.cala.fsu.edu/files/higher\\_order\\_thinking\\_skills.pdf](http://www.cala.fsu.edu/files/higher_order_thinking_skills.pdf)



- Klauer, K. J., & Phye, G. D. (2008). Inductive reasoning: a training approach. *Review of Educational Research*, 78, 85-123.
- Koretz, D. M. (2002, Autumn). Limitations in the use of achievement tests as measures of educators' productivity. *The Journal of Human Resources*. 37, 752-777.
- Koretz, D. M. (2005). Alignment, high stakes, and the inflation of test scores. *Yearbook of the National Society for the Study of Education*, 104, 99-118.
- Koretz, D. M. (2013). Adapting the practice of measurement to high-stakes contexts. Working paper, Harvard University, Cambridge, MA.
- Kumeh, T. (2011, March 25). Education: Standardized tests, explained. *Mother Jones*. Retrieved from <http://www.motherjones.com/mixed-media/2011/03/NCLB-standardized-tests-explained>
- Kuncel, N. R. & Hezlett, S. A. (2007). Standardized tests predict graduate students' success. *Science*, 315, 1080-1081.
- Kuzel, A. (1999). Sampling in qualitative inquiry. In B. Crabtree. & W. Miller (Eds.) *Doing qualitative research* (pp. 33-45). Thousand Oaks, CA: Sage.
- Meador, D. (2016, August 22). Examining the pros and cons of standardized testing. *ThoughtCo*. Retrieved from <https://www.thoughtco.com/examining-the-pros-and-cons-of-standardized-testing-3194596>
- McCrae, R. R. (1987). Creativity, divergent thinking, and openness to experience. *Journal of Personality and Social Psychology*, 52, 1258-1265.
- McNeil, L. M. (2000). *Contradictions of School Reform: Educational Costs of Standardized Testing*. Great Britain: Routledge.

- Minichiello, V., & Kottler, J. A. (2010). The personal nature of qualitative research. In V. Minichiello & J.A. Kottler (Eds.), *Your own personal dimensions* (pp. 7–8). Thousand Oaks, CA: Sage.
- Mujis, D. (2004). Introduction to quantitative research. In D. Mujis (Ed.) *Doing quantitative research in education with SPSS* (pp.1). Thousand Oaks, CA: Sage.
- National Council of Teachers of English. (2014). How standardized tests shape—and limit—student learning. Author. Retrieved from <http://www.ncte.org/library/NCTEFiles/Resources/Journals/CC/0242-nov2014/CC0242PolicyStandardized.pdf>
- Neal, D. (2013). *The consequences of using one assessment system to pursue two objectives*. Working paper, University of Chicago, Chicago, IL.
- Nelson, H. (2013). *Testing more, teaching less: What's America's obsession with student testing costs in money and lost instructional time*. New York: The American Federation of Teachers, 1-34.
- 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, 116-130.
- Phelps, R. (2002). Estimating the costs and benefits of educational testing programs. *Education Consumers Foundation*, 2, 1-3.
- Pintrich, P. R., & DeGroot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33-40.

- Polkinghorne, D. E. (2005). Language and meaning: Data collection in qualitative research. *Journal of Counseling Psychology*, 52, 137-145.
- Popham, W.J. (1999). Why standardized test scores don't measure educational quality. *Educational Leadership*, 56, 8–15.
- Ravitch, D. (2010). *The death and life of the great American school system: how testing and choice are undermining education*. New York, NY: Basic Books.
- Rhee, M. (2011, May 24). Accommodate don't discriminate. *Huffington Post*. Retrieved from [http://www.huffingtonpost.com/michelle-rhee/accommodate-dont-discriminate\\_b\\_866187.html](http://www.huffingtonpost.com/michelle-rhee/accommodate-dont-discriminate_b_866187.html)
- Robson, L. S., Shannon, H. S., Goldenhar, L. M., & Hale, A. R. (2001). *Guide to evaluating the effectiveness of strategies for preventing work injuries: How to show whether a safety intervention really works*. Cincinnati, OH: National Institute for Occupational Safety and Health Publications Dissemination.
- Standardized Test. (2015, November 12). In S. Abbott (Ed.), *The glossary of education reform*. Retrieved from <http://edglossary.org/standardized-test/>
- Standerfer, L. (2006). Before NCLB: The history of ESEA. *Principal Leadership*, 6, 26-27.
- Strauss, V. (2011, April 26). Unanswered questions about standardized tests. *The Washington Post*. Retrieved from [http://www.washingtonpost.com/blogs/answer-sheet/post/unanswered-questions-about-standardized-tests/2011/04/26/AFNRPlmE\\_blog.html](http://www.washingtonpost.com/blogs/answer-sheet/post/unanswered-questions-about-standardized-tests/2011/04/26/AFNRPlmE_blog.html)
- Sylwester, R. (1994). How emotions affect learning. *Educational Leadership*, 52, 60-65.

- Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics*. New York, NY: HarperCollins College Publishers.
- Testmasters. (2017). Testmasters *SAT classroom prep courses*. Retrieved from <http://www.testmasters.com/sat.html>
- Testmasters. (2017b). *Houston, Texas - SAT prep courses*. Retrieved from <http://www.testmasters.com/sat/texas/Houston/>
- Time Out From Testing. (2015). *Standardized tests: Terms & definitions*. Retrieved from <http://www.timeoutfromtesting.org/doesdefinitions.php>
- The Glossary of Education Reform. (2015). *Standardized test definition*. Retrieved from <http://edglossary.org/standardized-test/>
- The National Center for Education Statistics. (2012). The Institute of Education Sciences: IPEDS Data Center. Retrieved from <http://nces.ed.gov/ipeds/datacenter/>
- The Princeton Review. (2017). *SAT test preparation*. Retrieved from <http://www.princetonreview.com/college/sat-test-preparation.aspx>
- The Princeton Review. (2017b). *Guarantee terms and conditions*. Retrieved from <https://qa-www.princetonreview.com/legal/guarantee-better-scores>
- U.S. Congress, Office of Technology Assessment. (1992). *Testing in America's schools: Asking the right questions*. Washington, D.C.: U.S. Government Printing Office.
- U.S. Department of Education. (2003, June). *Welcome letter—no child left behind*. Retrieved from <http://www2.ed.gov/nclb/overview/welcome/index.html>
- Van Buskirk, E. (2009, December). Former test-prep exec plots industry makeover. *Wired Business*. Retrieved from <http://www.wired.com/business/2009/12/test-prep-internet/>

Weissglass, J. (1998, April 15). The SAT: public-spirited or preserving privilege?

*Education Week. 60.*

## **APPENDIX A: TRAIT DISCOVERY RESEARCH & INTERVIEW PROTOCOL**

In order to determine the characteristics that are present in students who achieve large score increases using test-preparation methods, we interviewed a number of individuals who work at test-preparation companies. For the purposes of this study, we believed that the ability to glean insight from individuals that goes beyond statistical data allows for a more descriptive, thorough analysis than would be captured by a quantitative analysis. Furthermore, these insights provided a real-world basis for creating items in a quantitative measure specifically designed for participants in test preparation interventions.

### **Research Design.**

In a preliminary research project, we were interested in determining instructor-identified personality traits that allow students to succeed using material taught in test-preparation courses. This portion of our study follows the five stages for critical qualitative research, the methodology discussed by Carspecken (1996) in his text, *Critical Ethnography in Educational Research*. In brief, this methodology allows a researcher to explore how individuals subjectively view a specific concept (a core tenet of qualitative research), and from their discourse to glean themes or trends. Information regarding participants and other design aspects of this project can be found in Appendix A. The interview protocol (Appendix B) used was designed based on Carspecken's (1996) methodology, which states that an interview protocol should consist of domains, or overarching topics of interest. For this study, the major domains included instructional practices, the purpose and structure of a test-preparation course, a comparison of preparation methods, and personality traits, or characteristics, present in students who experience large score increases with standardized test preparation. The interview

protocol was elaborated throughout every interview in an attempt to follow the thoughts of individual interviewees.

### **Participants**

Qualitative research is focused on describing human experiences and, therefore, may require the purposeful selection of participants who can provide the researcher with insightful information about the experiences being studied (Polkinghorne, 2005).

Interviewees for this study were test-preparation instructors who have worked in the test-preparation industry for at least three years, taught test-preparation courses for a variety of tests, taught at least 10 test-preparation courses, and have participated in standardized-test content development. Since standardized tests (STs) have similarities, instructors who teach more than one subject could provide a more thorough explanation of personality traits that affect ST score improvement than single-subject instructors could. Content development allows an individual to gain insight regarding the way in which ST questions are written, and offers a unique perspective not found in the classroom environment. Thus, instructors who have developed content for test-preparation materials will be able to answer questions regarding best practices from more than a teaching perspective.

Five instructors were selected to participate in this study, three women and two men, all of who met the inclusion criteria. All the female instructors have worked for a test-preparation company for over five years, have had some experience in the traditional K-12 classroom, teach test-preparation courses for numerous subjects, and have participated in content development. All of the female instructors hold master's degrees.

The two male instructors have worked in the test-preparation industry for four years



and had previously taught in a traditional classroom setting. Accordingly, both instructors were able to provide insight regarding the differences between traditional and non-traditional test-preparation instruction. One instructor holds a bachelor's degree, and one instructor holds a master's degrees.

### **Data Collection.**

Dialogical data generation (Appendix B and C) occurred using a series of five, semi-structured interviews with test-preparation instructors. As noted by DiCicco-Bloom and Crabtree (2006), an in-depth, semi-structured interview is often the primary data source for qualitative research, and such interviews allow the interviewer to probe into social and personal matters. Each interview lasted for approximately an hour and a half, and took place over Skype, an application that provides live-streaming video chat, due to the geographical locations of the participants. In addition to two audio recorders, video data was captured in order to view facial expressions and body language associated with interviewee responses. The multiple methods of data capture allowed for the triangulation of data. In addition to capturing the literal data, member checks were used to ensure that the findings aligned with the information conveyed in each interview.

**Data Analysis.** In order to ensure confidentiality, any identifiers were removed from the interviews, and pseudonyms were used in place of actual names and places. Thus, any and all documentation in this study uses pseudonyms as signifiers in the place of names. Upon conclusion of each interview, the interview was then transcribed, dissected, and coded until common themes from all instructors' experiences were identified.

**Findings.** Four common themes arose regarding characteristics present in students who achieve the greatest increases in standardized-test scores (STSs). All of the interviewed instructors felt that students who receive the greatest STS increases (1) have high belief in the value of standardized-test preparation, (2) maintain positive attitude toward material and the instructor, (3) are compliant and use the content-superseding techniques (CSTs), regardless of previous familiarity with the technique, and (4) possess an open mind and willingness to try alternative techniques. Based on these four themes, we developed a survey that measures the strength of these characteristics in individuals.

The following definitions of Task Value, Attitude, Compliance, and Open-Mindedness were determined after analyzing semi-structured interviews with test-preparation instructors and distilling the most important qualities for student score improvement. Task Value is high when a student does the following: (1) appreciates the worth of the task, regardless of whether or not it is interesting; and (2) acknowledges that understanding the task will serve a greater purpose or aid in future work. Attitude is high when a student does the following: (1) believes that the task is feasible and (2) approaches the task with a sense of confidence. Compliance is high when a student does the following: (1) adheres to task requirements, regardless of familiarity with method; (2) uses new methods to complete a task, regardless of comfort with method; and (3) strives to master new methods for task completion. Open-mindedness is high when a student does the following: (1) demonstrates willingness to attempt new methods; (2) analyzes whether or not a task can be solved in alternative ways; and (3) expresses a desire to expand knowledge.

## INTERVIEW PROTOCOL

### **Topic Domain One:** Instructional Practices

**Covert Categories:** As a lead-off domain, I am attempting to solicit information from the instructor regarding his/her position at The Princeton Review (TPR) standardized test (ST) preparation courses (PCs); background in ST prep; the instructor's personal teaching philosophy; the instructor's belief regarding the educational philosophy of a STPC.

**Lead-off Question:** How did you learn about TPR and what made you want to work for the company?

### **Follow-up Questions:**

1. What is your current position at TPR?
  - a. How long have you taught STPC?
  - b. Have you worked in other capacities at TPR?
  - c. How have your other roles complemented your role as an instructor?
2. Did you make an active choice to work for a test-preparation company instead of a traditional, academic institution?
  - a. If so, why?
3. In your opinion, what is the overarching strategy for improving STS?
  - a. Do you believe that this strategy is effective? Why or why not?
4. What is your teaching philosophy and how do you implement it in the classroom?
5. What is the educational philosophy of a STPC?
  - a. How does this philosophy align with your own?
  - b. How does it differ?

### **Topic Domain Two:** Purpose and Structure of STPC.

**Covert Categories:** The purposes of this category are to determine how STPC are taught; how and why, in the instructor's mind, a STPC can affect standardized test scores (STS); why a STPC is an effective means of preparation.

**Lead-off Question:** How does a STPC work?

**Follow-up Questions:**

1. How are STPC taught and what material is covered in a STPC?
  - a. Content and techniques
  - b. Teaching style
  - c. Student engagement
2. What information do students learn in a STPC that allow students to improve STS?
3. Do you believe that STPCs are an effective means of improving STS?
  - a. Why or why not?
4. What do you think makes STPCs effective?
  - a. How can students achieve large score increases over a short period of time?
5. Why do you think a STPC helps achieve increased STS?
  - a. Are there any particularly helpful aspects of a STPC?
6. How can a STPC help improve STS?
  - a. Consider the SAT and GRE, which both evaluate skills taught in high school. How can a STPC help students with these skills?
  - b. Why would a student need to attend a STPC instead of simply reviewing the material?

**Topic Domain Three:** Comparison of preparation methods.

**Covert Categories:** The purpose of this domain is to gain insight on the instructor's view regarding traditional, academic ST preparation versus STPC; how do traditional instruction and STPC differ; what is covered in a STPC that is not covered in the

classroom; how does teaching to the test in a STPC differ from teaching to the test in the classroom; why is either traditional or non-traditional preparation better.

**Lead-off Question:** In terms of potential benefit, do you believe that any ST preparation is good ST preparation? Why or why not?

**Follow-up Questions:**

1. What are the differences between traditional preparation and STPCs?
  - a. Content, techniques, structure, class size, teaching style
2. What are the similarities between traditional instruction and STPCs?
  - a. Content, techniques, structure, class size, teaching style
3. What do students learn in a STPC that they do not in the traditional classroom?
4. Do you believe that STs evaluate the information covered in traditional education?
  - a. Why or why not?
5. Could students learn the same material in a STPC that they could from any of the myriad of ST preparation books?
6. There is always discussion about teachers “teaching to the test” in the institutional setting. How is a STPC any different?
  - a. Is the content different?
  - b. Are there alternative strategies not learned in the classroom?
7. Do you think that students could achieve similar score increases with both traditional and STPC preparation?
  - a. Why or why not?
8. Do you believe that a STPC is more effective than preparing via traditional methods?
  - a. Why or why not?

**Topic Domain Four:** Student Characteristics

**Covert Categories:** The point of this domain is to determine the characteristics of students who do, or do not, benefit from STPC; has the instructor noticed specific student character traits in students who excel in ST prep courses; what do students who excel in

STPC do that others do not; does a student's learning style impact what the student gains from a STPC.

**Lead-off Question:** Do you believe that all students are capable of improving STS through STPC?

**Follow-up Questions:**

1. Consider all of the students whom you have taught over the course of your time at TPR. Have you ever had a student who exceeded your expectations, with regards to score improvement?
  - a. What did that student do to prepare?
    - i. Come to class, do homework, take practice tests, etc.?
  - b. How was your relationship with that student?
    - i. Did s/he feel comfortable communicating with you, asking questions, etc.?
  - c. What personality traits did the student possess?
    - i. Introverted vs. Extroverted
    - ii. Discrete vs. Abstract
    - iii. Logical vs. Emotional
    - iv. Structured vs. Adaptable
2. Have you ever encountered a student that, despite putting forth effort, was unable to improve STS after a STPC?
  - a. What did the student do to prepare?
  - b. Why do you think the student was unable to improve?
  - c. What did that student do to prepare?
  - d. How was your relationship with that student?
    - i. Did s/he feel comfortable communicating with you, asking questions, etc.?
  - e. What personality traits did the student possess?
    - i. Introverted vs. Extroverted
    - ii. Discrete vs. Abstract
    - iii. Logical vs. Emotional

iv. Structured vs. Adaptable

3. What do students who experience substantial score increases do that is different from students who experience moderate or no score increases?
  - a. Are there necessary steps a student must take in order to improve STS?
4. Is there a 'type' of student who is more likely to improve his/her STS?
  - a. What type of personality does that student have?
  - b. Does the student possess any specific character traits that are beneficial?
  - c. What does the ideal student do to prepare?
  - d. Are there any particular skills or traits that you believe are essential for improving STS?
5. Do you think that a student's preferred learning style affects his/her success in a STPC?
6. What, in your opinion, is the most important thing a student can do to achieve success in improving STS?

## **APPENDIX B: PORTION OF INTERVIEW ONE**



**20: 30**

- [1] **E:** So you bring up an interesting point about the fact that, obviously, all your students, or anyone's students are going to be unique in their abilities. With regards to standardized tests, do you think all students are capable of improving standardized test scores?
- [2] **S:** I think any student who wants to is capable of improving their score. Um. [OC: Interviewee sounds as if she is adamant that any student can improve scores.]
- [3] **S:** I know one of the teachers at my school last year had a quote on her wall that was like, "If you are determined to learn, no one can stop you. If you are unwilling to learn, no one can help you." And I totally believe that. If a student wants to improve on the test, they totally will.
- [4] **S:** Especially if they're into school. But if you are talking about high school or college, same thing. If they want to do well on something and want to improve on a standardized test, if they pay attention to what I teach them in a test-prep class, or do a lot of practice on their own, they can improve.
- [5] **S:** But I can't teach anyone who doesn't want to learn. Whatever, whatever it is; I can't teach them. [OC: Interviewee sounds as if she does not have patience for students who are unwilling to put forth effort.]
- [6] **E:** Right, right. Yeah. That would be extremely difficult. So, think back to all the students you've taught over the course of your time at The Princeton Review. Have you ever had a student who exceeded your expectations, based on the things you just mentioned, with regards to score improvement?
- [7] **S:** So, like a student who didn't try to hard, but still had major improvement? [OC: Interviewee sounds as if this is a silly question.]
- [8] **E:** I guess that is one way to interpret it! I was thinking, uh, a student who may have come in scoring about average, and just did everything you asked, and went up a tremendous amount.

[OC: Interviewer sounds caught off guard. However, she appears to recover and go with the flow.]

[9] **S:** Um. It's hard to say because most of my classroom students don't tell me how they did—even though I ask them to. So, I don't necessarily know... [OC: Interviewee sounds slightly annoyed, but used to the fact that, that students don't take diagnostic exams.]

[10] **E:** Ahhh...

[11] **S:** They are always telling us to email the students and be like, "Hey, how'd ya do? Did you get into those colleges? Let me know if you have any younger siblings that need tutoring."

[12] **S:** Just, to me, there is no way to get back into contact with an old student without it sounding like I'm trying to make money off of them or something. [OC: Interviewee sounds really irritated at the follow-up methods used by The Princeton Review.]

[13] **S:** So, you know, if it's a tutoring student, I'll text them and be like, "Hey, how'd ya do?" And most will get back to me. But, with the classes, they will occasionally email me and tell me what they got. But most of them don't.

[14] **S:** But I did have this tutoring student, it was only for six hours in January. She had taken the online course, but I don't think she did a lot of the homework, because I flipped through her book and was like, "You should have had that assigned and everything."

[15] **S:** But she was a really good student when she met with me. She did all the homework I asked her to do. She followed all the techniques—which is surprisingly rare—and she, like, remembered the things I taught her.

[16] **S:** She'd be like, "Oh Yeah, I did Plugging In," or whatever it is. She would right down the vocab, or write down her own word on the vocab problem like I told her to. [OC: Interviewee's tone sounds happy that a student was willing and able to use the techniques on standardized tests.]

[17] **E:** That's great!

[18] **S:** So yeah. She is going to get her scores back in a couple days. So, I don't know, hopefully, it will be a big improvement, since she didn't do that well the first time. [OC: Interviewee sounds hopeful regarding her students eventual test score results.]

[19] **E:** Did she take any diagnostics that you were able to measure? I mean, I guess it was only six hours. So.

[20] **S:** Yeah, so I had practice tests for her from her course, but I didn't have her take any with me.

[21] **E:** Gotcha. Um. Your students who do improve in your courses, um, do they have a significant score improvement?

[22] **S:** Um. Some of them do. But it's hard to tell, honestly, because a lot of them do not do any practice tests. I mean, I've had a few students who take all the required practice tests; I don't know, probably like five or six. [OC: Interviewee sounds slightly disappointed in the number of students who actually take diagnostic exams.]

[23] **S:** But most of them do only one. If they even do one. So it is hard to tell whether they improved.

[24] **E:** Right. Um. So students who do appear to be, like the student you were talking about who listens, and recognizes when to use the techniques, so on and so forth...what kind of personality traits do you think the students who experience score increases actually possess. So, for example, do you think they tend to be more introverted or more extroverted?

**20:35**

[25] **S:** In my experience, they tend to be more compliant, I'd say. Basically, you tell them to do something and they'll be like, "Ok, I'll do it. They are probably the kind of person who would do all the homework when the teacher assigns it, and all that kind of stuff. So, not necessarily, like, the smartest students, but the ones who will, like, say they are going to do something and they do

it. [OC: Interviewee sounds emphatic that intelligence is not the sole predictor of individuals who could benefit from a standardized test prep course.]

[26] E: So, follow-through?

[27] S: Yeah, probably, something like that.

[28] E: Ok. So, I'm going to give you a couple of sets of words, and if you could tell which you feel that students who experience score improvements possess, that would be helpful! So, do you think students who actually achieve score increases are more discrete or abstract thinkers?

[29] S: What do you mean by discrete? [OC: Interviewee sounds confused by the question posed.]

[30] E: Can they look at things in multiple ways...or do they just have to look at things in one certain technique.

[31] E: So, for example, if they are given a problem with variables in it, are they only going to do algebra, or are they willing to go for a more abstract method like Plugging In, which may not be the proper method. [OC: After a bit of silence, the Interviewer provides a question to clarify the issue at hand.]

[32] S: Usually the students who score better are the ones who are open to, or the ones who improve more, are open to using a new technique, because, um, as I said they are the more compliant students. [OC: The interviewee seems to be emphasizing the compliant nature of students who improve standardized-test scores through a prep course.]

[33] S: They are willing to trust me. I guess I could say trusting is another word for students who improve a lot. [OC: Compliant and trusting seem to be along the same lines.]

[34] E: Well that is definitely a good word!

[35] E: Do you find these students to be more logic-based or more emotion-based?

- [36] **S:** Um...probably more logic based because when I explain to them logically how it is a faster and easier way to do it, they are like, "Yeah, ok. That makes sense. I'll do it."
- [37] **E:** Ah, that's a really interesting way of looking at it. For some reason, I had been thinking the opposite, because I was thinking, well, if they are logical, they will just want to do it the mathematical way. But, I completely see what you are saying, because they are thinking of the logic of improving their scores—if that is what you're saying. [OC: Interviewer sounds as if she is pleasantly surprised by the interpretation of the interviewee.]
- [38] **S:** Yeah!
- [39] **E:** And then, do they tend to be more introspective or do they tend to be more extroverted?
- [40] **S:** Um...I'm not sure that it really matters whether they are more introverted or more extroverted. Or if it is something I really notice in my classes. Like, I don't know, I mean, with online classes I can't really tell because I just call on everybody.
- [41] **E:** Right (*laughter*)
- [42] **S:** So, sometimes you can get a sense by their tone about whether they are annoyed I called on them or not. (*laughter*) Um, but, so, it's not something I've really even identified with my students, I don't think.
- [43] **E:** That's totally fair! Um, do you believe there are any particular character traits and/or skills that are absolutely imperative that a student has to be successful for improving standardized test scores—other than trusting and willingness to do their work and follow through.
- [44] **S:** I mean, yeah, if they want to improve their scores, they just have to have the drive. If they want to like, like, do well in my class or whatever...they have to be compliant and follow...they have to trust I know what I'm doing. [OC: Interviewer seems to emphasize, again, the necessity of trust, compliance, and drive.]

[45] S: Like, I just started this big class a week ago, it's like a 20-person GRE class and one of my students during the break was, like, messaging me, or like, private chatting me, like, "Well. I don't think the answer to is right to this problem is right." And I was like, "No. I know it's right."

[46] *Both E & S laugh* [OC: Laughter appears as if, "What a silly student to question the teacher's authority.]"

**20:40**

[47] S: And he was really arguing with me! And I feel like people do that to me a lot. Well, not a lot, but like sometimes. Or I get the sense they don't like think I'm totally qualified because I look really young, I taught in elementary school, which people don't respect, and because I explain all the little steps. [OC: Interviewee seems somewhat irritated by the criteria on which people judge the validity of her teaching.]

[48] S: So that might come across as, like, "Oh yeah, I need to explain how to find one-fifth," but, like, to me, yeah, that's obvious. Like, yeah, I know how to find, how to solve fractions and stuff, but it is because I'm a teacher, and I know there are other people who don't know the stuff. [OC: Interviewer sounds like she recognizes the frustration of students who view her as incompetent, but will stick to her methods.]

[49] S: But, some people might take that as, you know, I don't know that much about math—but I actually do. So, um...

[50] E: Right.

[51] S: Yeah, so the students who don't think I know what I'm talking about, I don't know if they are going to trust that my methods are good. But, I appreciate that curriculum for The Princeton Review, typically for the first class, it kind of, like, tries to trick those students, so that they can see I know what I'm doing.

[52] E: Ah, right. (*a chortle*)

- [53] **S:** Cause that helps me a lot. I like it when I call on one of the high-scoring students and, this actually happened with that same student, like, I called on him to do another problem and, I was like, “Can you explain how you got that again?” And he was like, “Well, I did this and this and I got this.” Which was wrong. And I was like, “Yeah, could you just go through and explain that again?”
- [54] *Both S and E start laughing.* [OC: Both laugh as if it is somewhat gratifying to have overly confident students miss a question.]
- [55] **S:** And he was like, “Oh...wait.” So I was like, “Yesssss.”
- [56] **E:** Well played. *laughs lightly.* So it seems as if those who are more resistant and unwilling to adopt non-traditional techniques are those that will have less of a score increase.
- [57] **S:** Yes, but those are already the high-scorers anyway.
- [58] **S:** I don’t think I’ve had a low-scorer be like, “No, I want to do the algebra way.” They are always like, “Yes. Please. Tell me an easier way to do it. I love it. Just tell me and I’ll do it. I’ll do it.” [OC: It sounds as if low-scorers are much more willing to trust and comply to the techniques taught in a standardized-test prep course.]
- [59] **E:** Right.
- [60] **S:** Like, I don’t know. Usually those are the students that are only there for verbal. Um. So. One time I had an in-person class with only two people and one of them was a math major. So, I was like, wow, I’m sure she knows way more math than I do. Uh. Cause she got like a 170 or 169 on her practice test.
- [61] **E:** Nice!
- [62] **S:** Yeah, but, she was like really tolerant. Like, she was really nice. She never questioned me or, like, a lot of times with the high-scorers, when you ask them, “How did you do this problem on the homework,” they will explain the algebra way. I’m like, ok, but, “Just do it my way. Even if

you don't want to do it. Just to tell the rest of the class." And, like, she would always explain how to do it my way.

[63] E: Aw, that's nice!

[64] S: Yeah. Um. I mean, that's how it should be. I mean, I understand, like, if I were in a situation of the high-scorers, I would probably be the other kind of person who wants to show-off about how much I know and doesn't think the teacher knows enough math because it is some test-prep company. They are not like a professor or something. [OC: Interviewee sounds as if she empathizes with high-scorers.]

[65] E: Fair.

[66] S: So, I can relate to that point of view. But, having been trained, I do recognize, like...

[67] S: At first I was like, "Why would I plug in numbers? It's more fun to do it the algebra way!" But, once I learned how to do it and tried the strategies, I was like, "Yeah, that actually makes sense and it is an easier way to do it." [OC: Interviewee appears as if she initially thought the techniques were absurd. However, it appears as if she came to appreciate such techniques.]

[68] S: It doesn't make me think as much. Which. I enjoy thinking through the problem, but if you are just trying to do the questions quickly. Like, you could always go home and do more algebra, but for the test, just do the easy way.

[69] E: Because I am sure they are going to be going home and busting out the algebra... *laughter*

[70] S: They act like they really enjoy it, so I don't know.

[71] E: No, that's cool if they do!

[72] E: Do you think, because you've been talking about the difference between low-scorers and high-scorers, which I think is significant, do you think there's a specific range where these test-prep courses are beneficial and where they are not so beneficial?



[73] S: Yeah. I definitely do. If you are a high-scoring student, I don't think you are going to get that much out of the test-prep classes.

[74] E: What would you deem as a high-score, say on the SAT, out of 800?

[75] S: Oh, out of 800? Anything above 700. [OC: There is no hesitation; interviewee provides an emphatic answer.]

[76] E: Ok, I think that makes sense.

[77] S: Well, it just depends on what your weaknesses are though. If...it's kind of hard to say what somebody could improve on if they are already scoring in the 700s.

[78] S: Like, I think about, I didn't get a perfect math score when I took the SAT. I got a 760. I missed like two problems. But, like, it's not because I didn't have enough math knowledge or needed test-prep. It was because I wasn't paying attention or didn't read something correctly.

[79] S: I mean, if you are missing like three or four problems, it is hard to teach somebody how to be more careful on every problem. Or something like that. So.

**20:45**

[80] E: Right. So, I guess on the flip side, do you think it's not helpful for people scoring between 200-300?

[81] S: Yeah, that could be true because, um, you know, like, I have this GRE student right now, a tutoring student, who took the class with me, but he didn't really get much out of the class. Once we started doing tutoring, I was like, "Why don't you know this? I taught this to you!"

[82] S: But, like, even though I called on him the same as everyone else, and I actually, always thought he was on track. I never really realized how low he was until I started working with him

tutoring. So, you know, it's hard because some students are going to be more outspoken when they don't understand something.

[83] **E:** Mmmhmmm

[84] **S:** So, yeah, I actually think it's great when a student will say, "Hey, can you explain that again," or "What was the second step?" Yeah, that's great. Like, just ask and I'll explain it again. But some students don't ask and then I don't really know that they are not understanding something.

[85] **S:** Um...what was I going to say.

[86] **S:** Yeah, but it really depends on the teacher though. I really make a concerted effort to teach everything from the beginning because I've spent so much time working with, especially children, those really low in math. You just have to repeat everything over and over, you have to be really patient, and you have to anticipate what they are going to ask. And, so, to me, that comes really naturally now and, so, I'm always doing that kind of stuff. So, I feel if someone was taking a class with me and they are a low-scorer, hopefully they are still going to understand it because I try to teach it to a pretty low level. Um, or you know, at least making sure I explain every step.

[87] **E:** So, essentially what you are saying is, if it were just you and the student, and the student is starting, let's say at a worst-case scenario of 200. In that instance, if they actually listen and ask questions when they are confused, would you say that it could work?

[88] **S:** What could work?

[89] **E:** A standardized test prep-course...like a one-on-one tutoring.

[90] **S:** Well, it just depends on what you mean by work. I mean, can they achieve a 750? Probably not. [OC: Interviewer appears to note that there is a limit to improvement garnered by standardized test prep courses.]

[91] **E:** Right *laughter*

- [92] S: But, can they make a significant improvement? Absolutely. I mean, the lower your score, the more opportunity you have to improve.
- [93] S: So, um, I mean, if you are scoring that low, you probably just have really huge gaps in your content knowledge. So, that's the real reason a course isn't going to help...because they have to get that content knowledge first. So, you know, once they understand the content and the preliminary math fundamentals, then I think the test-prep strategies are going to be fine.
- [94] S: But, it's stuff like, I teach them how to plug in, but if they don't know how to find 20% of something, that's where they go wrong. So, it's not with the techniques, it's with the fundamentals.
- [95] E: Right. So, if you had to pinpoint an ideal initial scoring range for students to take standardized test prep courses, to gain the most improvement, what range do you think that would be?
- [96] S: Ummmm. Well, to get the most improvement, a lower score is better. *Laughter.*
- [97] E: I mean, mathematically, yes. *Laughter*
- [98] S: I dunno, yeah, to me, I'm like, "Give me the lowest student you have and I'll help them improve a lot!" [OC: Interviewee sounds as if it is much easier to improve low-scoring student scores than high-scoring student scores.]
- [99] E: Fair enough!
- [100] S: For a course, you have to come in with some content knowledge. Um, so...
- [101] E: Right. And you did say that anything above 700 is basically, not helpful.
- [102] S: Yeah, so, on the SAT scale, probably like a 500ish. That type of student would do best on a test-prep course.

[103] E: That makes sense since it's the average.

[104] S: Yeah.

**20:50**

## **APPENDIX C: PORTION OF INTERVIEW TWO**

— *Excerpt One* —

**9:30 PM**

- [1] **E:** What is the overarching strategy for improving STS?
- [2] **C:** The overarching strategy for improving standardized test scores is really what I was talking about earlier; understanding what the test is trying to do to you and how it is expecting you to react.
- [3] **C:** Because the thing about a standardized test is that the creators count on the idea that each answer choice will be picked by a certain number of people in order to create the normal distribution, on the bell curve, that they need. So, in order to be able to say, with confidence, that I expect this many people to choose A, this many people to choose B, this many people to choose C, or D, or E, as a test writer, you have to have a really clear idea about what would induce someone to pick that any of those answers.
- [4] **C:** What I've always believed as a standardized test prep instructor, is that if you can get people to understand why it is that the test thinks you will go for any given answer in any given situation, that's when you can change your behavior to respond to the test tricks, rather than simply reacting to the test tricks and getting it wrong.
- [5] **E:** So, are you saying that a...so, first, let me add, that's what you believe is the educational philosophy of a test prep course, such as those offered by The Princeton Review, Kaplan, Test Masters, or any of them?
- [6] **C:** Yeah. I believe if you can understand. Our pedagogical approach is that if you can understand what it is that the test is trying to do to you, you can avoid making the mistakes the test is trying to get you to make.
- [7] **E:** Definitely. I believe that that makes perfect sense. Um. Is that a different philosophy than what you had been taught in a traditional classroom...or anywhere?

- [8] **C:** In a school, or traditional classroom, and I assume that's what you mean by a traditional classroom, a regular school class.
- [9] **E:** Like a K-12 classroom, or anything like that.
- [10] **C:** Right. In a K-12 classroom, it's not about trying to trick people into picking the wrong answer. And that, I think, is the fundamental difference between what happens in a test prep classroom and what happens in a standard K-12 classroom. At school, teachers aren't interested in getting a certain percentage of the class to pick a wrong answer. Because a standardized test must create a bell curve with its results, the people who write a standardized test are very interested that a certain number of people pick a certain wrong answer.
- [11] **C:** Because a standardized test is very different from what the teacher at school is trying to do, it makes all the sense in the world that the pedagogical approach in a test prep classroom must (IC: emphatic!) be different than in a traditional classroom.
- [12] **E:** Absolutely, absolutely. Um, I completely agree. So, if you had to summarize...not summarize, but list, certain things that are learned in a standardized test prep course that are not learned in a K-12 classroom, what would those things be?
- [13] **C:** The biggest one that stands out is the approach to wrong answers. We spend a lot of time in a STPC teaching students how to recognize what isn't the answer, whereas in a traditional K-12 classroom, what teachers are focused on is imparting the right information and getting you generate the right information—whether that is picking the right answer on a test, writing the right things in an essay response.
- [14] **C:** What teachers are focused on is how to have students arrive at the right answer. What a test-prep course ends up being much more focused on is what a student can get rid of, and thus arrive at, the right answer. And, so really, the whole framework is split because the test is more about wrong answers than right answers. That to me is the largest difference.

- [15] **C:** Beyond that then, if you look at a test prep classroom, if a student sees a right triangle with legs three and four, we in the test prep industry are interested in training them to recognize that the completion of the pattern is 3-4-5; that the hypotenuse is 5.
- [16] **C:** In a traditional classroom, the teacher would be a lot more interested in making sure that students knew how to apply the Pythagorean theorem and understanding how the Pythagorean theorem actually worked. We in the test prep industry are more interested in, I don't care why it works, I just want you to recognize it, or for that matter, not fall for the other ones.
- [17] **C:** So, in a test prep classroom, there is a way in which it isn't about teaching the content most of the time, it's about teaching patterns in what the test is doing and how to respond to those.
- [18] **E:** So, just to, uh, confirm, the major difference between traditional and test prep courses would be the way in which you approach the problem. In test prep courses you are looking at it critically to find what is wrong with something, whereas K-12 courses are more concerned with the process in which you arrive...
- [19] **C:** ...At the right answer. Uh huh! Exactly! And because of the nature of that, a lot of the times, teaching students to avoid wrong answers has at least as much to do with pattern recognition, as it has to do with understanding content.
- [20] **E:** Right. So, would you say that pattern recognition is a necessary characteristic of students who succeed in STPC courses?
- [21] **C:** The ones who succeed the most? Absolutely.
- [22] **E:** Ok. Um. So this brings up an interesting question, since they seem that they are direct opposites in the way they approach a question. Do you think standardized tests actually evaluate content covered in traditional K-12 education?



[23] C: Um, I don't really believe they do an effective job of it, no. I think they purport to, but because of the way they are written, they can't really. The whole point of a standardized multiple-choice test has to do with the idea that a certain number of people have to fall for a certain number of wrong answers. So, it can't really.

[24] C: What the test is trying to do is to get certain people to miss a certain number of questions, whereas what you work on in school is, "Do you know the right answer?" So, I mean, you can find the right answer by finding the right answer, or you can find the right answer by getting rid of all of the wrong ones.

[25] E: Right.

[26] C: So, if what teachers teach you in school is to understand how things actually work, and the test goes way out of its way to trick you into thinking that something else works, then the two aren't really evaluating the same thing at all. The fact that the test is so actively trying to trick you, means it isn't evaluating what you know, but it is evaluating how easily it can talk you out of what you know.

[27] E: Right.

**9:40PM**

— Excerpt Two —

**9:45PM**

[28] E: This brings me to my next question of student characteristics. Do you believe that all students are capable of improving standardized test scores through a standardized test prep course?

[29] C: Well...

[30] E: And I know I used an extreme word of all, but I'm curious about your opinion.

- [31] C: Well, yeah, the issue here is all. There are always going to be students who, for whatever reason, it won't work for them. So, I've had students before with cognitive impairments or processing delays, where, in the time I had to work with them, they were not able to grasp what I was doing, and master it effectively.
- [32] C: Now, in some cases, I think that if I had had more time with those students, I could have gotten through to them in a different way, but I don't necessarily think that is true. So, when you are dealing with students who have cognitive impairments or disorders of that sort, you can't necessarily say that all students will improve with test prep.
- [33] C: However, if you've got a student who doesn't have any major cognitive impairments or anything like that, and the student will actually listen to me, do what I suggest, and practice what I am encouraging them to do, um, the scores go up.
- [34] C: Now, all those things have to be true. The student actually has to do what I say. The student has to actually try out and use the techniques. If the student doesn't use the techniques, I can stand there and teach at them until I'm blue in the face, but if they won't actually use the techniques, they won't improve their scores simply because I have articulated the techniques in their earshot.
- [35] E: Right. Um. So, considering everything you just said. You said they actually have to try the techniques, they actually have to listen to you, I completely understand where you are coming from with that. So, what personality characteristics, what does a person have to inherently have as a character trait in order to succeed?
- [36] C: Well, I might have an easier time answering the question in the other direction. Like, what character traits may stand in the way of students picking up on things effectively? Would it be ok if I approached it from that direction?
- [37] E: Uh, sure!

- [38] C: Ok. So. Um. Arrogance and a sense of self-importance reaaaally, reaaaally [*IC: emphasized this point a lot.*] cause problems. If you've got a student who decides, "Well, my way is the best way," and won't even try my way, despite, you know, the encouragement on my part, despite the fact that it worked for me, despite the fact that it is working for other students.
- [39] C: If a student is so convinced in his own mind that his way is best, that he'll just never try my way, and sometimes it is ridiculous. You can tell a kid, "Well, when you did it your way, you missed this question, this question, this question, and this question, but when you did it my way you got three questions right." You can show a kid all day long, but if he thinks his way is better, and is not willing to try it [*IC: a new way*], then he is never going to improve.
- [40] C: I've had plenty of students who, for instance, there scores were in the high 700s on whatever section of the SAT—it tends to happen with math, but it happens in other sections as well—but if you have a student who is scoring a 770 in math, and you try to suggest to him that, you know, using this technique, that isn't how you would have normally thought of it, but it'll help you clean up those 30 points you are missing, and he responds with, "No, my way is clearly best," [*IC: said in a goofy voice that implies a caveman attitude.*] then he is probably going to stay at a 770 and never actually get the last couple of points.
- [41] E: Right.

**9:50**

- [42] C: I know this because it is partly my own experience. When I was in high school, I just assumed that I knew best on how to handle things. It was really eye-opening when I went through Princeton Review training, realizing that these techniques, that we teach kids who can't do it the school way, it's so not even about that, it's about having a reliable way to arrive at an answer.

- [43] **C:** And, if a kid can accept that his own algebra is flawed, and I'll do it this other way—I'm going to plug in—then he can clean up those questions he is missing. But if he keeps insisting on doing algebra because, "Well, he just knows he is a good algebra student," *[IC: said in a goofy voice that implies a caveman attitude.]* then that kid is going to keep missing the questions that the test writers have engineered to be hard to get, even for good algebra students.
- [44] **C:** So, arrogance really seems to stand in the way of students getting to where they want to be.
- [45] **C:** Now, that doesn't necessarily mean that, in order to succeed, you need to be the most humble soul in all the world. I wouldn't say that humility is a characteristic of students who succeed. But, I will say that extreme arrogance stands in the way of succeeding. Does that make sense?
- [46] **E:** It does make sense. Um. And, so, along with this arrogance and, it kind of sounds like you are saying, stubbornness? Would that be fair?
- [47] **C:** Exactly. It's stubbornness versus flexibility. I often find myself telling students that the hallmark of a good test-taker is the flexibility in how you respond to questions.
- [48] **C:** So, the best test-taker has a couple of different tools in their toolbox. They know several different ways to approach a question, and they reach into their toolbox with, you know, a discerning eye and say, "This tool will be most appropriate in this case, this other tool will be most appropriate in this other case." So, it's about not being stubborn about just one approach, but rather recognizing that there are several ways to arrive at the right answer, and you should use whichever one is most appropriate for the situation.
- [49] **E:** Makes perfect sense! Would you say that would be open-mindedness?
- [50] **C:** Um, it's...I would say that is part of it, yeah. But it is not necessarily all of it.
- [51] **E:** Right.

- [52] **C:** Part of it is knowing that, like, because I have several options, let me make an educated choice about which option is best. So, open-mindedness is the first step to having multiple options. But, within that, being able to decide, “Ok, I’m going to do this one this way, I’m going to do this other one this other way.”
- [53] **E:** And, how does that relate to your relationship with the students? Like, what does the student have to...I suppose I’m asking, does the student have to trust you, have faith in you.
- [54] **C:** Yeah. There is a lot of that. And, it is one of the first things I take care of with every class that I teach. In the first ten minutes of the class, you know, in the first five minutes, you will hear me say, “I know this test is crappy and painful, but if you trust me, listen to what I have to say, try it out—you know, actually try to put these techniques into action—if you will do the things I tell you to do, I promise you, we can improve your score on this. And I say that to all of my classes within the first five or ten minutes because it is really important.
- [55] **C:** When I train other instructors, this too is a really important part of what I do. I had an experience recently where, as a trainer, I was called in to evaluate another teacher. I was asked to observe the teacher’s lesson, and at the end of the lesson, I had to make some recommendations about what the teacher had to do differently, or better. At the end of the observation, we actually ended up pulling the teacher’s certification. We said, you are busted back to trainee status, you cannot teach anymore right now until you go through training again.
- [56] **E:** Wow...
- [57] **C:** And she said, “But I’m doing all of the things I’m supposed to do.” And I said there is a difference between saying the right words and making students believe those words. She had other problems—she wasn’t even saying the right words—but she’d say, “Oh this is a good idea,” [IC: Said in a ditzy girl voice.] and I didn’t believe her because she wasn’t building trust, she wasn’t building buy-in, she wasn’t trying to present the ideas as, like,

compelling recommendations. It was just like, “Well, you could do this,” or “This is a thing...”

- [58] **C:** And, I can’t have that among the teachers in the classroom, simply because that trust is so important. If the students don’t trust the teacher, they are not going to try the techniques, and they are going to assume that their way is best.
- [59] **E:** Right...
- [60] **C:** You have to break that arrogance down! And, so, as a teacher, and as a trainer who is responsible for making sure our teachers are quality, that element of building trust, and getting students to trust you...you can’t do this job unless you get students to trust you and they are willing to do it.
- [61] **E:** Right. I’m curious now, what do you think the characteristics are of a good test-prep instructor?
- [62] **C:** A good test-prep instructor knows what the most effective way to do something is, knows several other ways to do the same thing, appreciates that those other ways are going to be more or less effective, and can convey to students in a compelling manner why those approaches are the right or wrong thing to do.
- [63] **C:** So, a teacher has to be able to think of something from several different perspectives. A teacher has to think of it from a test-writer’s perspective, as I talked about before, it is important that teachers have a clear sense of what the test is trying to do, so that they can teach their students how to respond accordingly.
- [64] **C:** But if the teacher, um, is also someone who knows several ways to approach something, and can break out whichever way is most appropriate to the student in that circumstance. You may find that you have a student where one explanation doesn’t click for them.

- [65] **C:** A good teacher can turn around and explain the same thing from a different perspective, take a different approach to it, and get that concept to click. And, sometimes, it is not the second explanation that makes the difference. Sometimes it is the third, or even the fourth, explanation that makes the difference. But a good teacher has multiple ways of approaching the same kind of issue, so that whatever way makes most sense to the student, the teacher can make that connection with the student.
- [66] **E:** No, that makes perfect sense. Um..
- [67] **C:** A good teacher, though, also has to be convincing. A good teacher has to build buy-in. And, part of building buy-in, is showing them why any particular way of doing it, is better than another way.

## **APPENDIX D: CONSENT FORM**



### **Consent to Participate in a Research Study: TACO Survey**

Eliz Markowitz of the University of Houston, Department of Curriculum and Instruction invites you to be a part of a research study that examines the personality profiles of future teachers and the way personality affects standardized test scores. The purpose of the study is to design better support programs for students who need to take TExES certification exams. We are asking you to participate because you are currently a student in the teacher certification program at The University of Houston.

If you agree to be part of the research study, you will be asked to complete the NEO-PIR personality inventory, the TACO survey about your behaviors as a student, a PPR pretest, and a PPR posttest. We expect the NEO-PIR to take approximately 45-60 minutes to complete, the TACO survey to take approximately 15 minutes to complete, and both the PPR pretest and posttest to take approximately 30 minutes to complete. Some of the survey questions ask about your emotions related to school and may be distressing to you as you think about your experiences. If you need to talk to someone about these feelings, please contact the University of Houston Counseling and Psychological Services (CAPS) by calling either (713) 743-5454 or visiting the CAPS office in room 226 of Student Service Center 1.

The primary benefit of participating in this study is a detailed, 7-8 page report regarding your NEO-PI-R inventory from Psychological Assessment Resources Inc. that will provide insight into your personality. Furthermore, we hope that this study will contribute to the improvement of support systems for those who are working to attain teacher certification in the state of Texas.

All data will be anonymized, and will not be able to link your survey responses to you, but they will know that you participated in the research because you will be asked to log in. The survey software keeps your identifying information separate from the answers you provide to the survey. We plan to publish the results of this study, but will not include any information that would identify you.

Participating in this study is completely voluntary and if your standing in the education program will not be affected should you choose not to participate. Even if you decide to participate now, you may change your mind and stop at any time. You may choose to not answer an individual question or you may skip any section of the survey. Your consent to participate will allow investigators to contact you in the future for further educational research. If you have questions about this research study, you can contact Eliz Markowitz, [emarkow@central.uh.edu](mailto:emarkow@central.uh.edu), Dr. Mimi Lee, [mlee@central.uh.edu](mailto:mlee@central.uh.edu), or Dr. Susan X Day, [sdlay@central.uh.edu](mailto:sdlay@central.uh.edu), University of Houston.

This project has been reviewed by the University of Houston Committee for the Protection of Human Subjects (713) 743-9204. If you have questions about your rights as a research participant, or wish to obtain information, ask questions or discuss any concerns about this study with someone other than the researcher, please contact the University of Houston Committees for the Protection of Human Subjects, [cphs@central.uh.edu](mailto:cphs@central.uh.edu).

By continuing these surveys, you are consenting to participate in this research survey.

## **APPENDIX E: ORIGINAL TACO SURVEY**

### TACO Survey

*Please read and answer the following questions. Select only one answer for each question.*

- 1) Please enter the last four digits of your PSID.

\_\_\_\_\_

- 2) Please select your gender.

Male      Female      Other

- 3) Please select your age.

18-21      22-30      31-40      41-50      51-60      Over 60

- 4) What is your highest level of education?

High School      Undergraduate Degree      Masters Degree      Doctorate

*Rate your level of agreement with each of the following questions, for which 1 indicates the statement is untrue of you and 7 indicates the statement is very true of you.*

#### Section One

- 1) In class, I prefer course material that really challenges me so I can learn new things.

1      2      3      4      5      6      7

Not at all like me      Neutral      Just like me

- 2) In class, I prefer course material that arouses my curiosity, even if it is difficult to learn.

1      2      3      4      5      6      7

Not at all like me      Neutral      Just like me

- 3) I think I will be able to use what I learn in one course in other courses.

1      2      3      4      5      6      7

Not at all like me      Neutral      Just like me

- 4) It is important for me to learn the course material in class.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 5) I think the course material in this class is useful for me to learn.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 6) Understanding the subject matter of this course is very important to me.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 7) I ask myself questions to make sure I understand the material I have been studying in this class.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 8) I ask the instructor to clarify concepts I don't understand well.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 9) I make good use of my study time for this course.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 10) I usually study in a place where I can concentrate on my course work.

1	2	3	4	5	6	7
1	2	3	4	5	6	7

Not at all like me                      Neutral                      Just like me

## Section Two

- 1) I am very interested in the content area of this course.

1	2	3	4	5	6	7
1	2	3	4	5	6	7

Not at all like me                      Neutral                      Just like me

- 2) If I study in appropriate ways, then I will be able to learn the material in this course.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 3) It is my own fault if I don't learn the material in this course.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 4) If I try hard enough, then I will understand the course material.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 5) If I don't understand the course material, it is because I didn't try hard enough.

1	2	3	4	5	6	7
1	2	3	4	5	6	7

Not at all like me                      Neutral                      Just like me

- 6) I'm certain I can understand the most difficult material presented in the readings for this course.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 7) I'm confident I can understand the basic concepts taught in this course.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 8) I'm confident I can do an excellent job on the assignments and tests in this course.

1	2	3	4	5	6	7
1	2	3	4	5	6	7

Not at all like me                      Neutral                      Just like me

- 9) I'm certain I can master the skills being taught in this class.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 10) When I study for this class, I set goals for myself in order to direct my activities in

each study period.

1	2	3	4	5	6	7
Not at all like me			Neutral		Just like me	

---

### Section Three

- 1) When I study for this class, I practice saying the material to myself over and over.

1	2	3	4	5	6	7
Not at all like me			Neutral		Just like me	

- 2) When studying for this class, I read my class notes and the course readings over and over again.

1	2	3	4	5	6	7
Not at all like me			Neutral		Just like me	

- 3) I memorize key words to remind me of important concepts in this class.

1	2	3	4	5	6	7
Not at all like me			Neutral		Just like me	

- 4) I make lists of important terms for this course and memorize the lists.

1	2	3	4	5	6	7
Not at all like me			Neutral		Just like me	

- 5) If I don't understand the course material, it is because I didn't try hard enough.

1	2	3	4	5	6	7
Not at all like me			Neutral		Just like me	

- 6) When I become confused about something I'm reading for this class, I go back and try to figure it out.

1	2	3	4	5	6	7
Not at all like me			Neutral		Just like me	

- 7) If I get confused taking notes in class, I make sure I sort it out afterwards.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 8) I make sure I keep up with the weekly readings and assignments for this course.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 9) I attend class regularly.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 10) Even when course materials are dull and uninteresting, I manage to keep working until I finish.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

## Section Four

- 1) I often find myself questioning things I hear or read in this course to decide if I find them convincing.

1	2	3	4	5	6	7
1	2	3	4	5	6	7

Not at all like me                      Neutral                      Just like me

- 2) When a theory, interpretation, or conclusion is presented in class or in the readings, I try to decide if there is good supporting evidence.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 3) Whenever I read or hear an assertion or conclusion in this class, I think about possible alternatives.

1	2	3	4	5	6	7
1	2	3	4	5	6	7

Not at all like me                      Neutral                      Just like me

- 4) I try to play around with ideas of my own related to what I am learning in this course.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 5) I try to relate ideas in this subject to those in other courses whenever possible.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 6) In a class like this, I prefer course material that really challenges me so I can learn new things.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 7) In a class like this, I prefer course material that arouses my curiosity, even if it is difficult to learn.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 8) I always try to understand what the teacher is saying, even if it doesn't make sense.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

Not at all like me                      Neutral                      Just like me

- 9) I prefer knowing many approaches to material than knowing a single approach.

1	2	3	4	5	6	7
1	2	3	4	5	6	7

Not at all like me                      Neutral                      Just like me

- 10) I generally find that there are multiple ways to approach a given task.

1	2	3	4	5	6	7
---	---	---	---	---	---	---

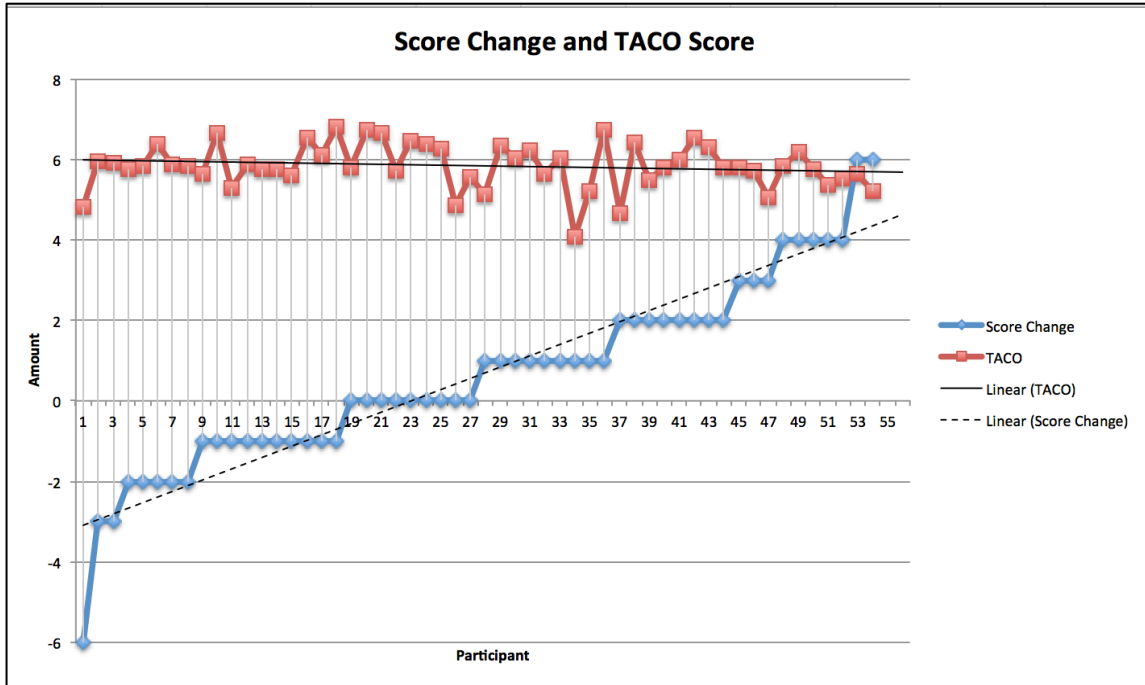
Not at all like me                      Neutral                      Just like me

**Thank you for your participation!**

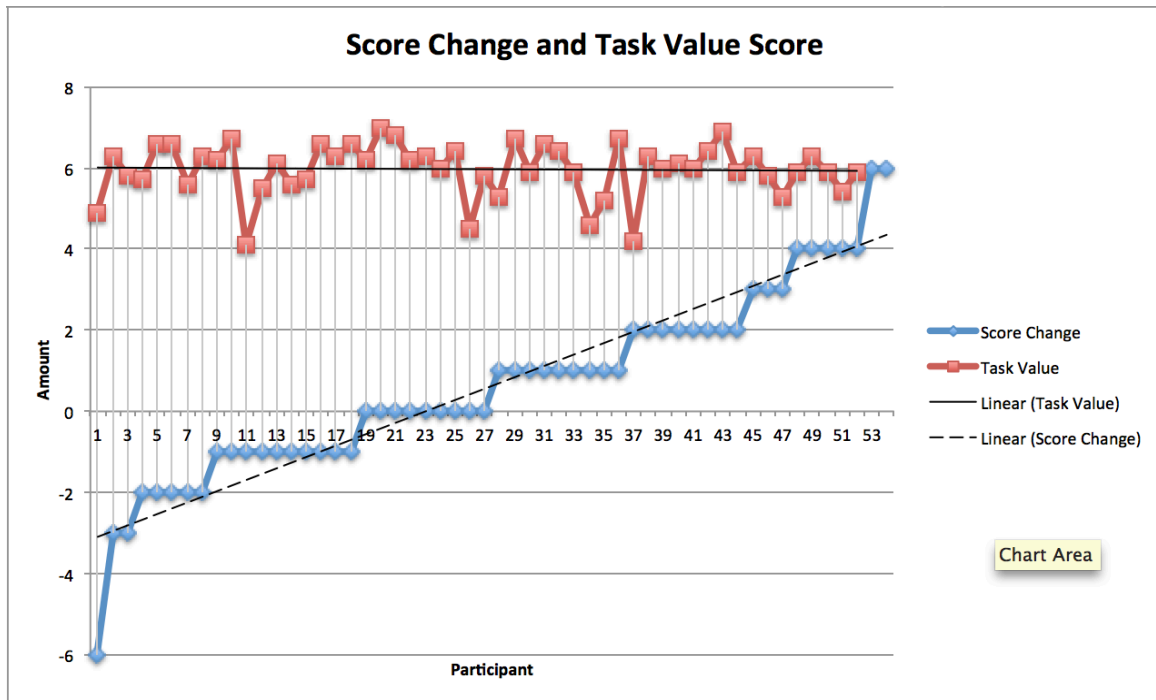


## **APPENDIX F: TACO PILOT STUDY RESULTS**

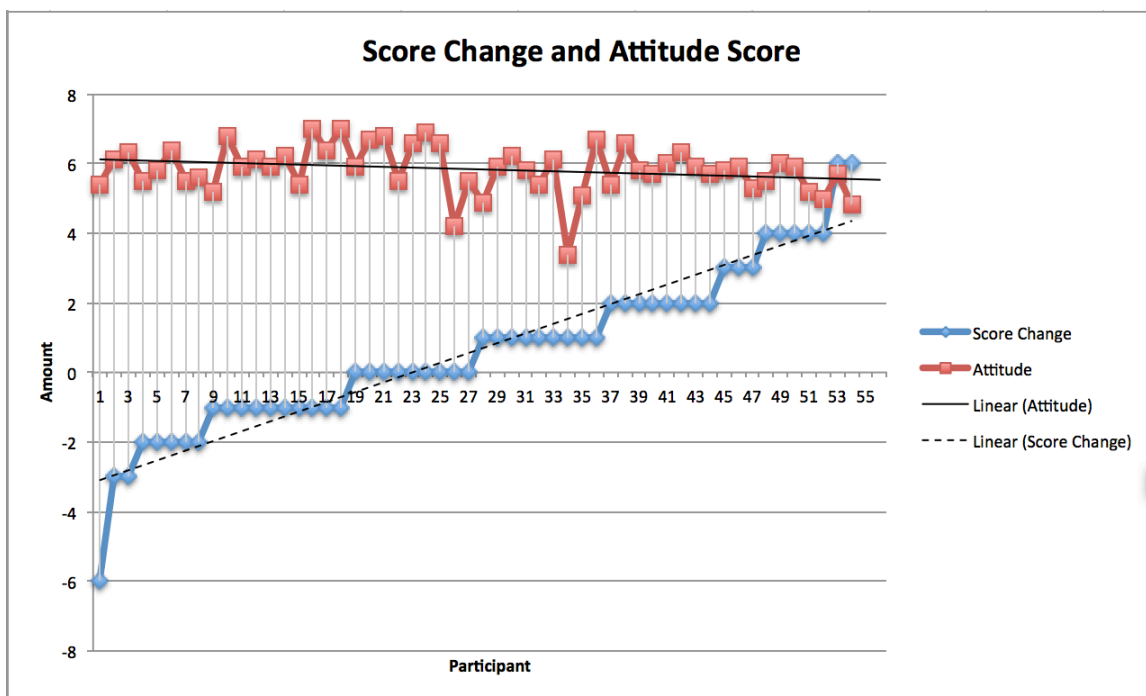
### Correlation between Score Change and TACO Score



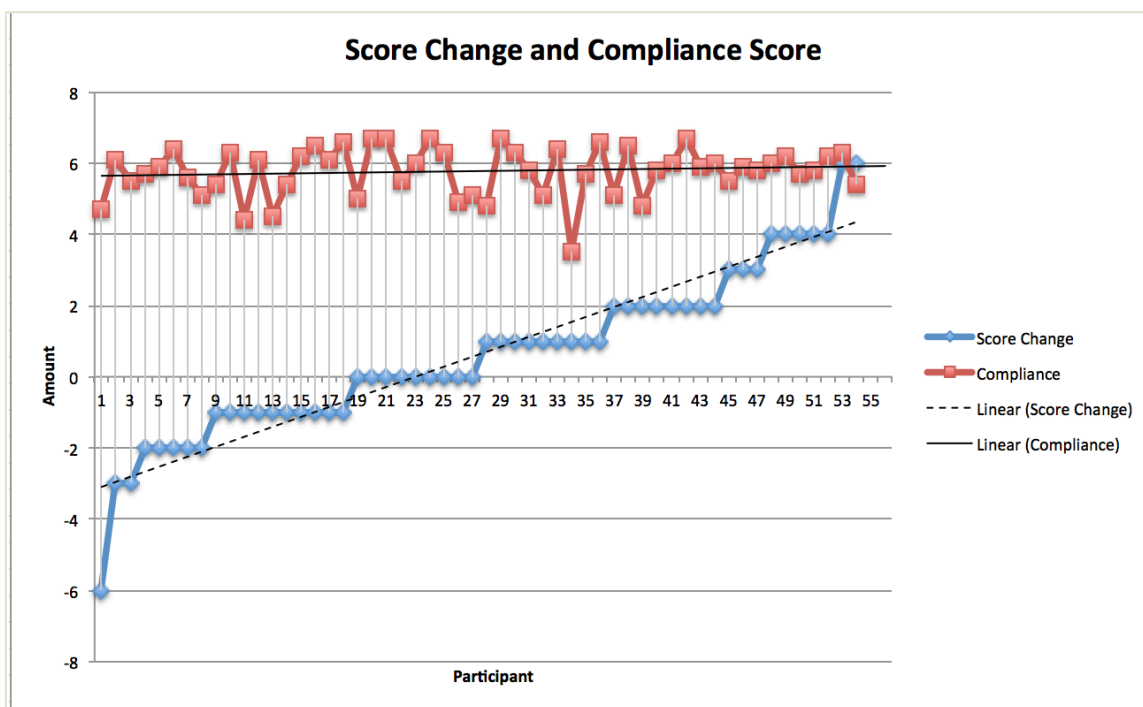
### Correlation between Score Change and Task Value



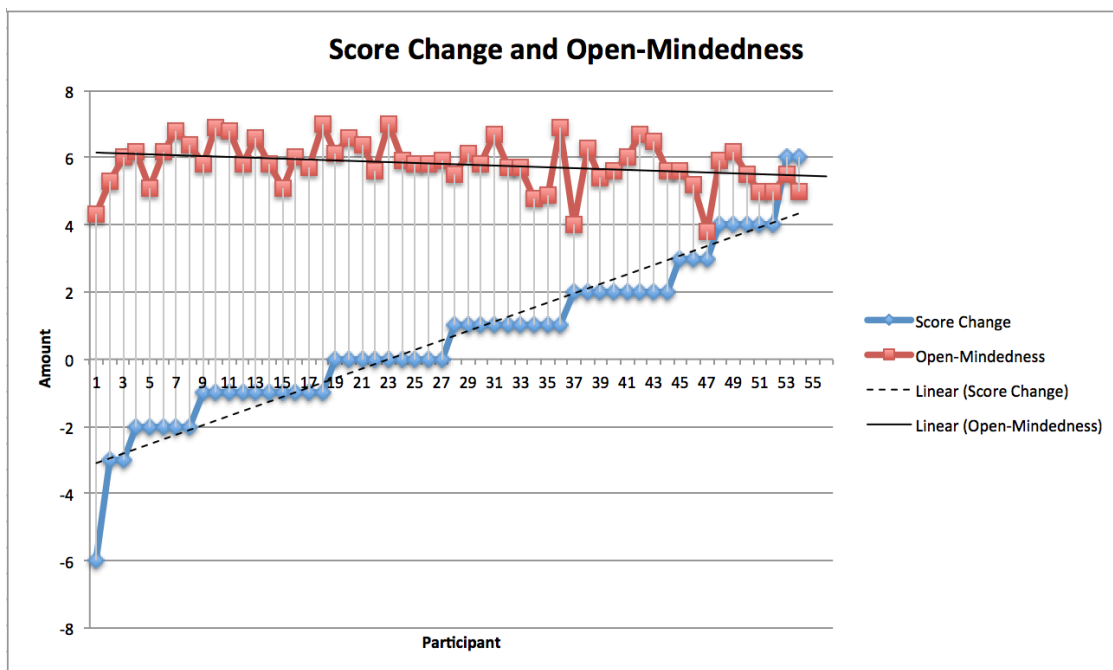
### Correlation between Score Change and Attitude



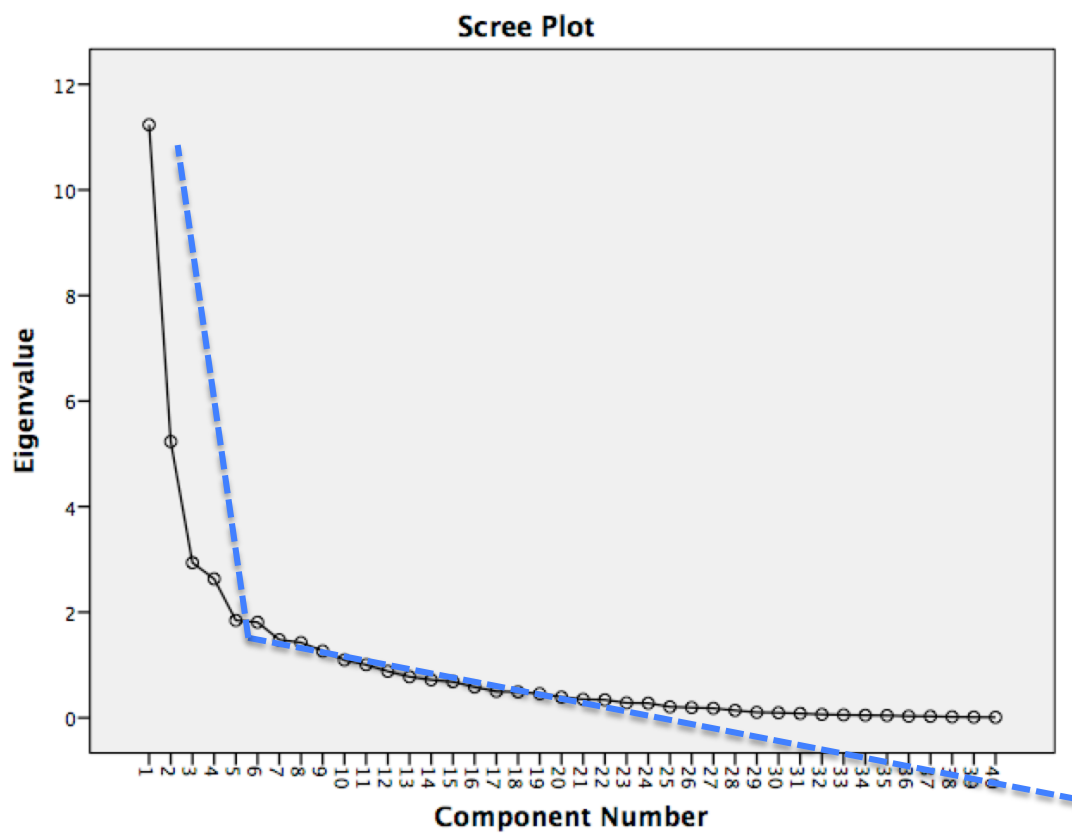
### Correlation between Score Change and Compliance



### Correlation between Score Change and Open-Mindedness



### Factor Analysis with Four Factors Extracted



Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	11.234	28.084	28.084	11.234	28.084	28.084
2	5.233	13.082	41.166	5.233	13.082	41.166
3	2.938	7.344	48.510	2.938	7.344	48.510
4	2.632	6.579	55.089	2.632	6.579	55.089

## **APPENDIX G: PPR PRETEST & POSTTEST**

*Please read and answer the following questions. Select only one answer for each question.*

1. To best ensure that all students in a fourth-grade inclusion classroom receive quality instruction, which of the following should lesson plans always include?
  - A) individual assignments for each students and planned time for whole-class reteaching
  - B) activities at varying difficulty levels to allow for students to make self-directed choices
  - C) extra time for reteaching and needs-based modifications of student activities
  - D) blocks of time for small-group instruction for students with learning disabilities
2. A third-grade teacher is planning a lesson on media literacy. One of the teacher's goals is for students to be able to compare various forms of written conventions used in digital media. Which of the following student activities best supports the goal?
  - A) Exploring a list of websites provided by the teacher and sorting the sites into categories based on the kind of information that is provided.
  - B) Researching a topic on the Internet and emailing a note to a friend that includes three facts from the research
  - C) Using a Venn diagram to distinguish characteristics of a blog from those in a Web-based newspaper article
  - D) Working in groups to create a classroom website about events in the classroom
3. A first-grade teacher has readers who are below grade level, each at a different level. In planning individualized instruction, which of the following is the best first step for the teacher to take?
  - A) Providing an assessment to identify students' reading strengths and weaknesses
  - B) Incorporating daily drills for students to practice reading skills
  - C) Having students pulled out of the classroom for small-group reading instruction
  - D) Developing students' oral language and writing skills
4. Which of the following teacher actions will most effectively promote a positive climate in an elementary classroom?
  - A) greeting students as they enter the classroom
  - B) setting the same goals for all students
  - C) allowing students to form their own groups during group work
  - D) providing lessons that are easy for students

5. When posing a question, the amount of wait time a teacher allows before calling on a student for a response primarily depends on the
- A) familiarity of the concept being discussed.
  - B) amount of time allotted for the lesson.
  - C) cognitive level of the question being asked.
  - D) number of students with their hands raised.
6. An eighth-grade science teacher decides to teach a complex unit in small, sequenced steps. Of the following, the major benefit of this approach is that it will:
- A) challenge students to independently synthesize the discrete concepts presented.
  - B) facilitate the teacher's creation of a valid and reliable end-of-unit assessment.
  - C) enable the teacher to target all instructional activities at a single conceptual level.
  - D) provide increased opportunities for students to process and reflect on the content presented.
7. As an informal assessment, a high school career and technology teacher has students complete an online job application and create an online résumé. The assessment is most beneficial because it allows students to
- A) apply their learning to real life
  - B) identify their vocational skills
  - C) apply metacognitive strategies
  - D) identify instructional goals
8. Ms. Nguyen notices that a boy in her kindergarten class comes to school very irritable and is hyperactive most of the day. He has trouble concentrating and frequently fights with other children. Which of the following actions would be the most appropriate first response to this situation?
- A) Referring the child to the school nurse for possible attention-deficit/hyperactivity disorder symptoms
  - B) Monitoring and documenting the behavior to be used at the parent-teacher conference in a few months
  - C) Setting up a conference with the child and parents to discuss behavior at home and sleep schedules
  - D) Collaborating with an administrator to develop an intervention plan to address behavior



9. Which of the following is the most effective way for a teacher to monitor student understanding of a task during group activities?
- A) Observing student interactions while circulating around the room
  - B) Remaining in an accessible location to answer student questions
  - C) Asking one member of the team to report on the group's activity
  - D) Providing a questionnaire for group members to evaluate one another
10. Ms. Lyons observes that a majority of the students in her class are not following a recently taught math process on how to regroup while subtracting. Which of the following would be the best way to reteach the math process?
- A) Selecting more complicated examples to better demonstrate the process
  - B) Connecting the underlying concepts to the process
  - C) Using abstract terms to explain the process
  - D) Presenting the process to a small group of students
11. Fifth-grade teachers at a school meet regularly to discuss strategies to help students struggling in mathematics. The primary benefit of the meetings is to help
- A) provide instruction for various types of learners
  - B) reward teachers who use effective teaching practices
  - C) acknowledge that some teachers use better methods than other teachers do
  - D) counsel each other when they are frustrated

*Use the pieces of information below to answer Questions 12-15. The pieces of information build on one another to create a hypothetical scenario, and the questions ask you to make appropriate decisions given the situation described.*

Mr. Hall, a middle school teacher, has the following goals while creating his behavior management program.

Goal 1: Develop an effective set of behavior expectations for students, including rewards and consequences.

Goal 2: Establish positive relationships with students' parents to get support for the behavior management program.

12. Which of the following will best help with effectively reaching the first goal?
- A) Posting a list of specific behavior standards on a poster at the front of the classroom
  - B) Allowing the students to discover the consequences of misbehavior through trial and error before discussing the importance of the behavior standards
  - C) Creating the behavior standards as a classroom community with student input and discussion
  - D) Inviting the school principal to participate during the discussion of the importance of the behavior standards
13. The most appropriate guideline to help Mr. Hall reach the first goal is to incorporate
- A) three to five general standards that guide productive learning.
  - B) ten to fifteen standards that provide detailed information.
  - C) three to five standards that define consequences.
  - D) ten to fifteen standards that teachers implement across classes.
14. Which of the following could Mr. Hall do to best achieve his main goal related to parents?
- A) Meet with the parents as soon as the students violate the behavior standards
  - B) Call the parents when help is needed to maintain the standards appropriately
  - C) Send an introductory letter home on the first day of school that asks for a conference in the first month
  - D) Solicit parental involvement when there is an after-school grade level activity
15. After the plan has been implemented, Mr. Hall schedules a phone conference with the parents of a student who is not following the plan. In the conference, which of the following approaches will best promote the second goal?
- A) Developing behavior standards to be implemented at home and then summarizing how to address academics and behavior
  - B) Preparing notes on the student's misconduct and then suggesting ways for the parents to provide positive and negative consequences
  - C) Providing a more specific set of behavior standards and then suggesting that the parents email any questions and/or concerns
  - D) Summarizing the student's achievements and challenges and then encouraging the parents to discuss the situation

## **APPENDIX H: REVISED TACO SURVEY**

## TACO Assessment

PSID: \_\_\_\_\_

Age: \_\_\_\_\_

Gender: M F O

Certification Level: EC-6 MS HS Non-Cert

Mark your level of agreement with each of the following statements by placing an "X" in the appropriate box; "SD" indicates that you strongly disagree, "D" indicates that you disagree, "N" indicates that you neither agree nor disagree, "A" indicates that you agree, and "SA" indicates that you strongly agree with the statement.

#	Statement	SD	D	N	A	SA
1	I think I will be able to use what I learn in one course in other courses.					
2	If I try hard enough, then I will understand the course material.					
3	I memorize key words to remind me of important concepts in class.					
4	In classes, I prefer course material that arouses my curiosity, even if it is difficult to learn.					
5	It is not important for me to learn the course material in class.					
6	It is my own fault if I do not learn the material in a course.					
7	For a course, I do not make lists of important terms and memorize the lists.					
8	I do not think about the possible alternatives when I read or hear an assertion or conclusion in a class.					
9	I think the course material in class is useful for me to learn.					
10	If I do not understand the course material, it is because I did not try hard enough.					
11	I make sure I keep up with the weekly readings and assignments for a course.					
12	In classes, I prefer course material that does not really challenge me.					
13	Understanding the subject matter of a course is not very important to me.					
14	I am certain I can understand the most difficult material presented in the readings for a course.					
15	When I become confused about something I am reading for a class, I go back and try to figure it out.					
16	I generally find that there are multiple ways to approach a given task.					
17	I usually study in a place where I can concentrate on my course work.					
18	It is not my fault if I do not learn the material in a course.					
19	I do not keep up with the readings and assignments in courses.					
20	When a theory, interpretation, or conclusion is presented in a class or in readings, I try to decide if there is good supporting evidence.					

Office Use Only:

T: \_\_\_\_\_ A: \_\_\_\_\_ C: \_\_\_\_\_ O: \_\_\_\_\_