

Spring 5-18-2012

An Investigation of the Combined Assessments Used as Entrance Criteria for a Gifted English Middle School Program

Colleen D. Lindsey Ph.D.
University of New Orleans

Follow this and additional works at: <http://scholarworks.uno.edu/td>



Part of the [Gifted Education Commons](#)

Recommended Citation

Lindsey, Colleen D. Ph.D., "An Investigation of the Combined Assessments Used as Entrance Criteria for a Gifted English Middle School Program" (2012). *University of New Orleans Theses and Dissertations*. 1459.
<http://scholarworks.uno.edu/td/1459>

This Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UNO. It has been accepted for inclusion in University of New Orleans Theses and Dissertations by an authorized administrator of ScholarWorks@UNO. The author is solely responsible for ensuring compliance with copyright. For more information, please contact scholarworks@uno.edu.

An Investigation of the Combined Assessments Used as Entrance Criteria for a Gifted English
Middle School Program

A Dissertation

Submitted to the Graduate Faculty of the
University of New Orleans
in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy
in
Special Education and Habilitative Services
with an emphasis in Gifted Education

by

Colleen D. Lindsey

B.S. Baylor University, 1991

M.S. Baylor University, 1992

May, 2012

Copyright 2012, Colleen D. Lindsey

Dedication

I dedicate this to my parents, Rick and Priscilla Donowho. Plain and simple, you are the only reason I was able to consider beginning, continuing, or finishing this degree. You are the most amazing parents anyone could have; I know I am blessed beyond measure to call you mine. For the past fourteen years – with a seven year hiatus in the middle – you have supported, loved, believed, and encouraged me to make this dream a reality. If that weren't enough, you extended that support in innumerable ways during these last few years as I have written this dissertation. You have cared for our children, cleaned the house, washed countless loads of laundry, paid tuition bills, run errands, and so much more. I suppose there is one richer blessing than having been raised by you two; you are instrumental in the lives of our children. You are continuing to pass on your thirst for knowledge and guiding them toward such magnificent futures. I dedicate this to honor you.

Acknowledgments

First and foremost, I give the greatest thanks to God. I am convinced that He orchestrated so many factors to make this possible. Only by relying on Jesus' infinite grace and endless providence, was I able to complete this degree program. I give Him all the glory and I pray that I continue to walk in the path He has laid out for me.

Next, I must express my deepest gratitude and appreciation for my family's support. My husband's confidence in my ability to complete my dissertation gave me the encouragement necessary to persevere. I would never have had the determination on my own to continue through the hardships. He is my best friend, and I am blessed to walk this life holding his hand. I am also grateful that my children willingly allowed me to work on many weekends. They turned down the television in the next room, gave me space when I was working, and encouraged me along; their excitement to return for my graduation kept me going at times. It certainly has added to my own momentum in this last sprint to the finish. Katie, Justin, and Meredith you are my shining stars, my source of inspiration, and my best work to date. I hope my perseverance reminds you to pursue your goals, even when it is difficult or feels impossible. Lastly, I want acknowledge my parents' support. It always extended beyond what I could have expected or asked for; it truly made this all possible. Thank you is such a trite and tired expression; it cannot begin to convey the depth of my gratitude and appreciation to my family. I love you all!

Thank you to the members of my committee from The University of New Orleans: Dr. Mary Cronin, Dr. Paul Bole, and Dr. Kate Reynolds. I also want to thank Dr. Susan Johnsen of Baylor University and Dr. Jerry Keating of the Statistical Consulting Center at The University of Texas at San Antonio, directed by Dr. Stephanie Cano. Each made significant contributions to

my doctoral program and dissertation. None of them began this journey with me, but all were willing to join my committee to help me finish. I appreciate the countless hours and willingness to see this to the end. The numerous emails, phone conference calls, and meetings to revise and edit this document solidified a true scholar-in-the-making. I also want to acknowledge the foundation laid by Dr. Mary Banbury and Dr. Bud Wellington. Without Dr. Banbury's relentless encouragement, I might not have begun this program at all. Bud, I finally jumped through this last hoop and finished ... as a passionate advocate for gifted students everywhere! Because of your combined efforts, I know I can defend my theory and my practice of gifted education.

I also need to acknowledge the influence of my cousin, Jack Smith. When I finished my Master's program and began teaching, he told me that I would never return to school to finish my Ph.D. He assured me that marriage and children would supersede any of my personal goals. While I did take a hiatus because of my family, I am pretty proud of my ultimate accomplishment and wouldn't change a thing. Oddly, I owe him for fueling the fire at times; proving him wrong has been quite satisfying.

Finally, I have learned that advanced degrees take a village to complete, and I would be remiss to not mention a few of those who contributed to my success. Thank you Kay Stotts, Lynn Belt, Kendall Biehl, Jennifer Wilson, Kelly Bigley, Louise Menchaca, Kim Hawkins, and countless others who helped me collect data, listen to me as I talked endlessly about this process, explained statistics I had long forgotten, and were willing to provide insight when needed. You were part of my village, and I am eternally grateful.

Table of Contents

List of Tables	x
Abstract	xi
Chapter One: Introduction	1
Purpose of the Study	2
Theoretical Framework.....	3
Background of the Study	5
Intelligence.....	7
Aptitude and achievement tests	7
Gifted education.....	9
Local programming.....	10
Need for the Study	12
Research Questions.....	13
Null Hypothesis	13
Limitations and Assumptions of the Study.....	14
Limitations	14
Assumptions.....	16
Definitions.....	17
Organization of the Study	28
Summary	28
Chapter Two: Review of the Literature	30
Components of Study.....	31
Purpose of Study	32
Program and identification agreement	32
Measurement of success	33

Evaluation of entrance criteria	34
Parallel Studies.....	34
Progressive Achievement Test.....	36
Scholastic Aptitude Test	36
Controversy.....	37
Kindergarten Diagnostic Instrument.....	39
Math and science.....	40
Intelligence Theorists.....	42
Psychometric approach	42
Multiple forms of intelligence	43
Aptitude and Achievement Tests	45
Identification of gifted learners.....	45
Psychometric approach	46
Multiple forms of intelligence	48
Controversy.....	50
Gifted Education	52
National.....	52
State and local education agency (LEA).....	56
Special Population	57
English language learners (ELL)	59
Minorities.....	60
At risk students	62
Girls.....	64
Four Tests Used in Local School District.....	66
Cognitive Abilities Test (CogAT)	66
Iowa Test of Basic Skills (ITBS).....	67
STEP Writing.....	68

Naglieri Nonverbal Ability Test (NNAT).....	70
Summary.....	72
Chapter Three: Methodology and Procedures	73
Purpose of the Study	73
Research Design.....	74
Research Questions.....	75
Null Hypothesis	76
Participant Selection	76
Procedures.....	79
Data Collection	80
Instruments.....	81
Cognitive Abilities Test (CogAT)	81
Iowa Test of Basic Skills (ITBS).....	82
STEP Writing.....	83
Naglieri Nonverbal Abilities Test (NNAT)	84
Data Analysis	85
Limitations	85
Summary	87
Chapter Four: Results	89
Demographic Description	90
Methods and Procedures	92
Research Question #1	95
Research Question #2	97
Research Question #3	99
Chapter Five: Conclusions.....	105
Summary of the Findings.....	105
Findings.....	106

Discussion	107
Aptitude and achievement.....	107
Grades	108
Teacher effectiveness.....	109
Gender.....	110
Programming match.....	110
Individual students	111
Implications for Gifted Programs	112
Recommendations for Further Study	115
Summary	117
References.....	119
Appendices.....	154
Appendix A: Middle School English Identification Matrix 6-8	154
Appendix B: GT Teacher Checklist – Secondary English.....	155
Appendix C: IRB Exemption Letter	156
Appendix D: NEISD Permission Letter.....	157
Appendix E: NEISD Gifted Director Permission Letter.....	158
Appendix F: Campus Principal Permission Letter.....	159
Vita.....	160

List of Tables

Table 1: Demographic Information and Participation Years Display	91
Table 2: Total Sample Summary	93
Table 3: Prediction of Success	94
Table 4: Prediction of Success Coefficients	94
Table 5: Four Entrance Assessments Correlation.....	96
Table 6: Regression Analysis: Reading and Step Significance	97
Table 7: Logistic Regression: Reading and Step Significance	98
Table 8: Repeated Measures ANOVA Without ID: Reading and Step Significance	100
Table 9: Nine-Week Grades and Averages.....	101
Table 10: Repeated Measures ANOVA With ID: Individual Student Significance.....	102
Table 11: Logistic Regression: Gender.....	103
Table 12: Linear Estimation: Prediction of Success.....	104

Abstract

The purpose of this study was to determine if the four assessments for entrance into an academic middle school gifted English program were accurately predicting success, as measured by students' grades each nine-week grading period. Some students were dismissed from the program each year because they could not maintain the required minimum average of 80%. The four entrance assessments evaluated were the Cognitive Abilities Test (CogAT), Naglieri Nonverbal Ability Test (NNAT), STEP Writing test, and the Iowa Test of Basic Skills: Reading (ITBS). The sample consisted of 150 sixth, seventh, and eighth grade students studied longitudinally over the span of four years from a suburban middle school in a large Texas school district.

Using correlation, logistic regression, and generalized linear regression models, the results showed that all the students selected to participate in the middle school gifted English program were statistically capable of success, whether they successfully remained in the program or not. Additionally, the results indicated that the two achievement tests (ITBS Reading and STEP Writing test) better predicted which students were successful, whereas the aptitude tests (CogAT and NNAT) did not. The achievement tests were determined to be better predictors of students' success, as measured by grades, in this rigorous academic middle school gifted English program. Other findings include (a) students' grades increased over time in the program, (b) females were predicted to earn about two point higher grades than males, and (c) the individual student was a significant predictor of success based on entrance scores.

Finally, several recommendations were made for future research. These possibilities include repeating this study using standard scores for data analysis rather than the percentile

scores that were available for this investigation. An additional recommendation is to investigate a possible replacement for the STEP Writing test, as it has not been nationally normed in decades. Another possibility would be to evaluate the curriculum and teacher effectiveness within the district using the NAGC (2010b) Pre-K-Grade 12 Gifted Programming Standards. A final potential study could implement specific interventions for use with students at-risk for underachievement to determine which strategies are most effective.

gifted education, assessment, academic gifted program, repeated measures, Texas

Chapter One

Introduction

The struggle to identify and serve academically gifted students has been a prominent topic of discussion in education for about a hundred years (Assouline, 1997; Jolly & Bruno, 2010). It remains a highly debated topic due to the range of theories about how to accurately define and identify giftedness as well as how to educate gifted individuals (Johnsen, 1997; Parker, Jordan, Kirk, Aspiranti, & Bain, 2010; Reis & Renzulli, 2010; Subotnik, Olszewski-Kubilius, & Worrell, 2011). Most recently, as school budgets have tightened, evaluating a gifted program can be part of the “key to the survival” of such programs (Johnsen, 2011, p. 151). Schools throughout the nation have struggled to find effective tools to identify students for gifted programs. This study investigated the relationship between the entrance criteria and student performance in a Texas middle school gifted English language arts program.

The district policy, beginning in the spring of 2007, was to evaluate all previously identified gifted students as they concluded fifth grade to determine the continued appropriateness for placement in the middle school gifted programs, and to evaluate students referred by their teacher or parent. The policy was changed because once admitted, some of the students did not find success in the gifted middle school English course and were removed. However, removal from the gifted program can be detrimental to both the student and the integrity of the program (Matthews & Shaunessy, 2010; Van Tassel-Baska, 2006). It calls to question whether the students were identified correctly, whether the program was constructed based on sound criteria, or if other intervening factors (e.g. student motivation, life events,

teacher effectiveness, curriculum appropriateness) were to blame (Lohman, 2005a; Louis, Subotnik, Breland, & Lewis, 2000; Nettelbeck & Wilson, 2005; Samuels, 2007; Sternberg & Clinkenbeard, 1995). Evaluating a gifted program is essential (Dai, Swanson, & Cheng, 2011; Doina, 1997; Hunsaker, 2000; Mathews & Burns, 1992; Tomlinson & Callahan, 1994; Van Tassel-Baska, 2006); this study attempted to evaluate the entrance criteria component.

The intent of this study was to evaluate the efficacy of the combined evaluation instruments used to identify students and those students' success in the gifted English program in a large Texas school district (65,000 students). As of November 2011, this district had 65 campuses including 12 magnet programs. The district student demographics are approximately African American 7.0%, Asian 3.5%, Caucasian 31.1%, Hispanic 55.1%, and other 3.2%. Success was determined by the grades students earned in the gifted middle school English class, as that is the criteria the school district uses to dismiss unsuccessful students.

This chapter discusses the background of the study, the purpose and theoretical framework of the study, the need for the study, as well as the research questions and limitations of the study. It concludes with the organization of the study and the definition of terms.

Purpose of the Study

The purpose of this study was to compare the entrance assessments used to determine student eligibility using the grades each student earned. Students are held accountable for their performance once admitted to the gifted English middle school program. Students with grades less than 80% for two consecutive nine week grading periods are dismissed from the program. The aim of the study was to build a foundation upon which the entrance criteria can be evaluated

for this specific district's middle school gifted English program, determining if all four instruments are the best evaluation tool to identify students for the gifted English middle school program.

Theoretical Framework

Educators' fascination with intelligence has been a relatively recent focus of investigation. Aside from the advanced studies of isolated prodigies in the past centuries (Jolly, 2004), this field has been explored in earnest for only about a hundred or so years (Kane & Brand, 2006). From the beginnings, when Sir Francis Galton published *Hereditary Genius* (1869) and attempted to differentiate between the mental abilities of various individuals in 1885 (Thurstone, 1946), to Binet, Simon, and Terman, the empirical data and theories regarding intelligence has grown exponentially (Assouline, 1997; McClellan, 1985; Parker et al., 2010; Reis & Renzulli, 2010; Sisk, 1990; Thurstone).

Numerous theorists have attempted to define intelligence (Gottfredson, 2004). The range of models is extensive. Some theorists, like Spearman, view intelligence as a "general, unified capacity for acquiring knowledge, reasoning, and solving problems that is demonstrated in different ways..." (Weinberg, 1989, p. 98). Other theorists, like Guilford and Thurstone, hold the view that intelligence is composed of many separate mental abilities (Weinberg). Yet others, like Horn (1982) and Vernon (1940), hold intermediate positions between the first two groups. A final group includes theorists who view intelligence as manifested in multiple independent domains, like Gardner (1983), or that intelligence includes analytic, creative and practical abilities, like Sternberg (1985).

Some theorists have developed surveys or alternate assessment instruments such as Sternberg's Multidimensional Abilities Test (Sternberg, 1985) instead of using more traditional psychometric tests (e.g., Cognitive Abilities Test [Lohman & Hagen, 2001]; Wechsler Intelligence Scale for Children [Wechsler, 2003]; Stanford Binet Intelligence Scales [Roid, 2003]). However, traditional psychometric tests are typically used to evaluate and inform professionals about a child's capabilities (Nettelbeck & Wilson, 2005). The field has been inundated with traditional psychometric tests because they are useful in educational settings for a variety of needs (Nettelbeck & Wilson).

Assouline (1997) pointed out that measuring intelligence is merely a matter of semantics and that intelligence test scores are not accurate representations of intelligence. However, she claims, "a well-designed test can provide relevant and useful information about behaviors that are typically associated with intelligence" (p. 91). Most theorists recognize that all cognitive tests measure developed abilities in some way (Lohman, 2006).

In this endless sea of intelligence tests, this Texas school district's gifted education director selected four instruments to serve as a gatekeeper for the middle school gifted English program. Relying on the results of traditional psychometric tests, students are evaluated for inclusion in the middle school gifted English program. These tests measure cognitive abilities and achievements in both reading and writing. Because of the number of students dismissed from the program each year (ranging from 2%-8%), a closer examination was warranted. Based on evaluating the relationship between student semester grades in the gifted English course and these entrance measures, recommendations for successful identification and placement were given to the district.

Background of the Study

Beginning with the 2006-2007 school year, the school district changed its policy regarding the academically gifted middle school program. Students enrolled in the elementary gifted program did not automatically move to the middle school gifted classes as they had before. The district began re-testing gifted students as they finished their fifth-grade year, to ensure that all gifted students were qualified for the rigorous and demanding high school curriculum taught in the gifted middle school English and mathematics programs. The sixth-grade students that entered the gifted English middle school program for the 2007-2008 school year were the first to have undergone this reevaluation process. School district-mandated longitudinal data collection began with this group of students as they entered sixth grade.

The implied expectation was that only capable students would enter the academically gifted middle school English program. In a perfect world, no gifted student would be missed (incorrectly unidentified), and no gifted student would be forced to exit the program because the course was too difficult for them. However, students continued to be removed from the program.

It is unacceptable to blindly assume the entrance criteria are appropriate (Hansen & Linden, 1990; Lind, 2001; Lohman, 2005a; Louis et al., 2000; Sternberg & Clinkenbeard, 1995). Gifted students already face challenges (e.g. perfectionism, depression, underachievement, overexcitabilities, social issues) (Cross, 2009b; Cross & Coleman, 1993; Delisle, 1986; Delisle & Galbraith, 2002; Grobman, 2006; Lind; Morisano & Shore, 2010; Peterson, 2009). The misuse of these evaluation instruments can erroneously place students in the academic gifted program, where they struggle and may ultimately be dismissed due to insufficient grades.

An additional concern developed because some students, identified as gifted in elementary school, are told during the reevaluation process they are no longer qualified to continue in the academic gifted program. Such news is a crushing blow to many adolescents (Adelson, 2007; Cross, 2009b). Their self-esteem is potentially shattered (Berlin, 2009). Some gifted students struggle with perfectionism, underachievement, and other social-emotional difficulties, so the inaccurate assessment data may hinder their future academic success (Delisle, 1986; Hargrove, 2008; Hollingworth, 1938; Peterson, 2006). There are potentially dozens of other reasons why students might not perform as expected in the academically gifted English course (e.g., motivation, life events, teacher effectiveness, curriculum appropriateness), but a reasonable first step was to evaluate the efficacy of the entrance assessments before other criteria was researched.

A final concern develops when the integrity of the program is questioned as students are dismissed. Are the entrance criteria appropriately identifying students for the gifted program? Expecting zero student removals is unrealistic as there are other factors that affect student performance. However, a reasonable first step in this process was to evaluate and determine if the entrance criteria were effective (Matthews & Shaunessy, 2010; Van Tassel-Baska, 2006).

The following sections explore several elements of the background of the study. The first three expand on the theoretical base for intelligence, aptitude and achievement tests, and gifted education. The last section details the history of the local programming in the Texas school district.

Intelligence.

Many argue about the definition of intelligence and how it should be measured. Thorndike in the 1920s and Thurstone in the 1930s supported a narrow definition of innate mental capacity, which was further researched and supported by Spearman (1925). However, Witte argued that a high intelligence in childhood did not determine genius in adulthood (Jolly & Bruno, 2010), and more recently Gardner's *Frames of Mind* (1983) and Sternberg's Triarchic Theory (1985) are well known for challenging the general mental ability (Spearman's *g*) as the only definition of intelligence (Nettelbeck & Wilson, 2005). Theorists like Renzulli, Dweck and others have also challenged the narrow definition of intelligence supported by Spearman and his followers (Dweck, 2010; Weinberg, 1989).

The gifted middle school English program in this study is based on scholastic aptitude and achievement. Other programs (e.g., Schoolwide Enrichment Model) might use a broader definition of intelligence and include creativity or motivation to qualify students. Since the program being studied is academic in nature and based on scholastic aptitude and achievement, using a broader definition (e.g., including creativity or kinesthetic intelligence) would not be appropriate (Delisle, 2003; Louis et al., 2000).

Aptitude and achievement tests.

Regardless of the debate, theorists recognize that some students learn differently and at a faster rate than the average learner, and instruments to measure their potential and success vary widely (Gordon, Duff, Davidson, & Whitaker, 2010; Johnsen, 1997). Aptitude tests (e.g., IQ) have been used in the United States since the Binet-Simon scale was translated at Stanford

University in 1916 by Terman and his colleagues. According to Leslie (2000), they “adapted the content for U.S. schools, set new age norms and standardized the distribution of scores so that the mean score would always be 100” (p. 3), and its uses became broader (e.g., vocational placement, evaluate mentally retarded students, identify learning disabilities, academically gifted learners) as the Stanford-Binet Intelligence Scale. From the time of Terman’s work (e.g., 1920, 1921, 1925, 1939) until today, the methods for determining intelligence are as diverse as the varied definitions of intelligence (Johnsen). Additionally, the methods for measuring achievement are as seemingly endless as the subject matter that can be measured by an assessment.

Many theorists argue that achievement and aptitude measure the same concept; that semantics is all that separates the two constructs (Lohman, 2006). This potential interrelatedness can cause concern when both aptitude and achievement measures are used to make judgments about students. If the two assessments are strongly interrelated, then a student would essentially be judged twice on the same ability if both aptitude and achievement were measured and presumed to be separate indicators. However, some students demonstrate a relatively high aptitude - but have quite low performance (Peterson & Shinn, 2002). This discrepancy was the legal definition of a learning disability in the 1977 Individuals With Disabilities Act (U.S. Department Of Education [USDOE], 1977). For many years, this discrepancy was the foundation for identifying students with specific learning disabilities (Büttner & Hasselhorn, 2011; Clark, 1997; Maehler, & Schuchardt, 2011; McKinzie, 2009; Peterson & Shinn; Pyryt, 1996). While the discrepancy model has been criticized and replaced by the Response to Intervention (RtI) Model, as supported by Stanovich (2005) and others, it does not negate the fact that students can show significant discrepancies between their measured aptitude and achievement. There are

differences between what an aptitude and achievement tests measure (Assouline, 1997). Many academically gifted programs, including the one in this study, use the combination of aptitude and achievement scores to determine eligibility for prospective students.

Gifted education.

The federal definition serves as a basis for some states' definitions, but it is not required. Additionally, "there is no federal policy that oversees how states should handle gifted education" nor mandates regarding funding (Samuels, 2007, p. 2). Each school district is given the opportunity to interpret its state's definition and programming options for gifted students. Consequently, there is great diversity within the field of gifted education throughout the nation (National Association for Gifted Children [NAGC], 2009). Since the local district has not published a definition of giftedness, this study will defer to the state definition. The state of Texas currently defines giftedness as:

"Gifted and talented students" means a child or youth who performs at or shows the potential for performing at a remarkably high level of accomplishment when compared to others of the same age, experience, or environment and who: 1) exhibits high performance capability in an intellectual, creative, or artistic area; 2) possesses an unusual capacity for leadership; or 3) excels in a specific academic field. (Texas Education Code, 2009)

One of the challenges in developing healthy gifted programming is selecting appropriate identification measures, which must be conducted by qualified personnel. As consumers of test data, directors of gifted programming, at the state and district levels, must be well educated

regarding the effectiveness and purposes of the measures they utilize. “Identification of gifted children ideally is tied to the theory of giftedness to which one subscribes. Pragmatically, however, it more often is tied to the particular program for which one is seeking participants” (Louis et al., 2000, p. 295). Louis and colleagues emphasized that programs ought to select students who will profit from placement in the program, but that sometimes excludes at-risk or culturally diverse populations. Fletcher, Denton, and Francis (2005) further noted that the measures used to identify students must be evaluated for their reliability and validity. That is, assessment instruments should be evaluated to determine if they are effective measures so other issues can be addressed and researched. Of utmost concern are the students who are identified as academically gifted - but do not find success in the academically gifted educational coursework. Whether or not the student is a poor match for the program, struggling with underachievement, or facing difficult life events, being dismissed from a gifted program is a difficult transition. Those kinds of disappointments are great (Cross, 2009a; Hansen & Linden, 1990; Hollingworth, 1938). Gifted students often struggle with social and emotional issues; programmatic flaws should be evaluated and addressed (Delisle, 1986; Hargrove, 2008; Peterson, 2006). There must be careful consideration of the type of programming and curriculum for students prior to testing and placement (Louis et al., 2000).

Local programming.

The local education agency (LEA) in this study began its program for academically gifted children in 1978. This Texas school district began with an elementary level program pull-out enrichment program – as it is at the present time. Changes in the middle and high school

programs did not appear in the LEA until 1982, when gifted students were provided with segregated, homogenous classes in English and mathematics.

As of 2007, previously identified gifted elementary students moving to middle school are given a set of assessments to evaluate their capabilities and determine continued eligibility for the gifted programs. Students are evaluated for the middle school gifted English program using the Cognitive Abilities Test (CogAT) (Lohman & Hagen, 2001) and the Naglieri Nonverbal Abilities Test (NNAT) (Naglieri, 2003) as measures of aptitude, as well as the reading portion of the Iowa Test of Basic Skills (ITBS) (Hoover, Dunbar, & Frisbie, 2001), and the STEP Writing test (Educational Testing Service, 1957).

Each student's combined assessment scores are evaluated, and their placement is based upon their performance using a matrix (Appendix A) developed by the LEA gifted director. In order to qualify for the middle school gifted English program, students must score at least a 95 percentile on the majority of the assessments. As expected, students sometimes have two scores above that range and two below. In those cases, a committee of qualified gifted education teachers evaluates the students' performance, as well as the course averages, performance on state tests, teacher checklist of observed gifted behaviors (Appendix B), and any anecdotal information (e.g., parental anecdotal data, extracurricular activities) available to evaluate the student. No single test score can keep a student out of the gifted education program. This is because the program director relies on a combination of scores to fairly evaluate students for admission to the gifted program.

Once enrolled in the gifted program in middle school, each student is required to maintain an 80% average in their gifted classes each nine-week grading period to remain in the

program. If a student's average falls below 80% for any given grading period, they are placed on probation. Students can be dismissed from the gifted program if they fail to earn 80% or above in the subsequent grading period. While Sternberg (2008) and others may argue that "education needs to move away from single targeted measures of success, such as grade point average (GPA) (p. 152)," it is important to realize that this is currently how this LEA chooses to measure success in the program. The purpose of this study was to compare the efficacy of the entrance assessments with the grades students earn because these elements are what the LEA selected as criteria for all students to enter and remain in the academically gifted English middle school program.

Need for the Study

Most students in the academic gifted English program are successful. However, 3% during the 2007-2008, 8.3% in the 2008-2009, and 2% during the 2009-2010 school years were dismissed due to grades below the 80% cut off for two consecutive nine-week grading periods. Was there a problem with the measures used to identify students for the gifted education program initially or other intervening factors? This dissertation evaluated the efficacy of the combined assessment measures of the four entry instruments and the relationship with the grades that students earn in the gifted English class. "This level of accountability not only improves the overall program but also helps the public see how the program is effectively identifying students who need services in gifted education" (Johnsen, 2011a, p. 152). Findings from this dissertation were used to make recommendations to the LEA. Alternatively, there is evident need for future focus on other possible intervening factors (e.g., curriculum appropriateness, teacher effectiveness, motivation).

Research Questions

The purpose of this study was to determine whether the combined scores on four aptitude and achievement measures correlated with student grades for students enrolled in the middle school gifted English program in a particular Texas school district. The following research questions were designed to structure the study:

1. Is there a correlation between the percentile scores earned on the battery of tests given and the middle school gifted English student grades?
2. To what extent did the battery of tests used as entrance criteria for the middle school gifted English program positively correlate with student grades?
3. Which entrance criteria most strongly predict academic success in the middle school gifted English program?

Null Hypotheses

A null hypothesis cannot be proven, but a set of data can reject or fail to reject. The lack of statistical difference may not indicate that there is no difference in reality, but that there is not enough evidence to reject the null hypothesis. The null hypotheses for this study state that there will be no correlation between the evaluation measures used to qualify students and student grades for a middle school gifted English program, and that the scores on the combined entrance tests do not relate to student grades. Additionally, higher scores on the entrance measures do not relate to higher student grades. Lastly, no single entrance measure predicts academic success in the middle school English program better than the combination of the four instruments.

Limitations and Assumptions of the Study

There were several limitations to this study, weaknesses that potentially limit the reliability and validity of the results. These include limitations associated with the eligibility assessments used and with other student factors. Some statistical limitations also apply. Additionally, numerous assumptions were made in this study. These assumed truths were centered on the teacher, qualifications of school personnel, appropriateness of curriculum, assessments, and the LEA practices for removal.

Limitations.

The first limitations were centered on the aptitude and achievement tests used to identify students for the academically gifted English middle school course. The study was limited by intervening factors affecting students during most test administrations (e.g., life events like lack of sleep the night before testing, distractions in the testing environment, or troublesome morning or home environment).

Among the tests used was the STEP Writing test. Limitations with this assessment include its older publication date and lack of recent national norming. Robins and Jolly (2011) recommend national norming samples more than 14 years old be used with “great caution” because of the changing national demographics (p. 81). Additional limitations associated with the STEP Writing test have to do with the changes in writing foci and instruction as our society has evolved (e.g., addition of technology and related terms, variance of expectations over time regarding style, changes in grammar, and spelling rules). The methods and instruction used for teaching writing continues to change as educators prepare students for our ever-changing world.

Just as the definition of literacy has changed (Ntiri, 2009), so does the definition of appropriate writing and its instruction. Technology, and its associated words and specific ways of communicating, has been incorporated into our writing and its instruction. Additionally, the emphasis on narrative writing versus expository writing has changed over time. These changes may or may not be accurately evaluated with the STEP Writing test. The choice to use this instrument was made by the district gifted education coordinator, not the researcher of this study. Indeed, this and several other limitations may become the focus of a follow-up study, which will be further discussed in Chapter Five.

Other limitations are associated with students. One limitation is related to newly immigrated students regarding their language acquisition, reading skills, and test-taking fears, as these limit the accurate results of many aptitude tests (Lohman, 2006). A further student limitation relates to non-identified academically gifted students who were never recommended for evaluation or who do not score well enough on the assessments to participate in the academically gifted English middle school program.

Another serious concern is the student grades that are used as a mark of success in the gifted program. Limitations associated with student grades include teacher subjectivity, students' personal interest in the subject matter, teacher instructional effectiveness, and motivation, to name a few (Stanley & Baines, 2004).

Several statistical limitations also apply. One is the limited range of scores inherent in the data. This lack of variability is a concern because the population was strictly limited, since all of the students included in the study had passed the 95 percentile cut-off for admittance into the gifted program. For most subjects, a majority of their scores on the four entrance assessments

were high. Additionally, since most of these students had previously demonstrated high achievement, it was expected they would perform well in gifted coursework. However, because the range of scores was limited by the number of students recommended for evaluation initially, and the range of grades was limited to averages of 80 or above in most cases, the correlation or variance was relatively low (Pyryt, 1996). Keeping these statistical limitations in mind, further investigation is warranted since some students did not perform well once admitted to the gifted academic program.

Assumptions.

This study was based on several assumptions. The first of these were that teachers who are assigned to teach gifted education courses had appropriate certification and training to serve this population, had an interest in and good rapport with the gifted students, had adequate training in their content field, used strategies to identify students at risk for underachievement then created appropriate intervention plans, and used a wide variety of differentiated instruction methods to meet the needs of the diverse population typically found in academically gifted classrooms. A related assumption was that limiting the study to students taught by a single teacher will eliminate the variable of student grades, since “grades are not comparable across teachers” (Robins & Jolly, 2011, p.82).

A second set of assumptions were related to the decision-making policies regarding the entrance assessments and student inclusion. The first assumption was that the district personnel that selected the entrance assessments and had re-normed the STEP Writing test were qualified to do so according to appropriate ethical and procedural guidelines. A related assumption was that when students have earned two scores above 95 percentile and two scores below 95

percentile on their entrance tests, the personnel involved in making a decision regarding placement for these children were qualified to make those kinds of determinations. A final assumption was that the LEA's procedure and protocol for removing unsuccessful students from the gifted English middle school program was appropriate.

A third set of assumptions were related to the curriculum used in the gifted English middle school program. An assumption was that the district-created curriculum was appropriate for the age of these identified gifted students. Furthermore, an assumption was made that the teacher was using this curriculum as it was designed to instruct and assess the students.

Finally, a fourth set of assumptions related to the aptitude and achievement tests. A first assumption for this study was that aptitude and achievement are interrelated, but independent measures (Assouline, 1997). Related assumptions were that the aptitude and achievement tests used were the most appropriate for the audience and setting, as well as the assumption that the definition of giftedness being used by the district was appropriate.

Definitions

For the purposes of this study, the following definitions were used to ensure consistency:

Acceleration

While interpreted in several ways, acceleration is an educational intervention decidedly effective for high-ability students. A student completes material intended for an older audience (e.g., first grade student who completes math with the second grade class) (Colangelo et al., 2010).

Achievement

Something accomplished, especially by superior ability or special effort. Level at which a student can demonstrate acquired knowledge in a specific domain or field (e.g., mathematical computation skills, phoneme awareness, identification of parts of a frog) (Colangelo & Davis, 1997).

Achievement Tests

Achievement tests are assessments that are designed to determine the degree of knowledge and proficiency exhibited by an individual in a specific area or set of areas. An achievement test is sometimes administered as part of the acceptance process into an educational program. It is an instrument used to measure a specific domain or combination of domains. Discrepancy between aptitude and achievement scores sometimes indicates learning disabilities or other school related issues (Francis, Fletcher, Stuebing, & Lyon, 2005).

Aptitude Tests

A standardized test designed to predict an individual's ability to learn certain skills. It is also called an IQ or intelligence test. The average score is 100. Often used in school settings to determine giftedness and other exceptionalities (Assoluline, 1997).

At-risk Students

The at-risk indicator codes from the Texas Education Agency (2010) give 13 different criteria for identifying a student as being at-risk for dropping out of school. The 13 indicators are listed on page 62.

Attention Deficit (Hyperactivity) Disorder (ADD or ADHD)

ADHD is a lifelong disorder that affects brain and results in a variety of behaviors including lack of self-control and impulsivity. Individuals must exhibit symptoms for at least six months and be severe enough to impair the normal development of the individual (Schlachter, 2008; Taylor & Larson, 1998).

Cognitive Abilities

See Intelligence/Aptitude

Creativity

Refers to the phenomenon whereby a person creates something new (a product, a solution, a work of art etc.) that has some kind of value. What counts as "new" may be in reference to the individual creator, or to the society or domain within which the novelty occurs. What counts as "valuable" is similarly defined in a variety of ways.

Scholarly interest in creativity ranges widely: the relationship between creativity and general intelligence; the mental and neurological processes associated with creative activity; personality type and creative ability; creativity and mental health; creativity in education; and ways of fostering creativity through training and technology (Clark, 1997; Colangelo & Davis, 1997).

May be used as a component for determining giftedness (Gagne, 2004; Renzulli, 1986).

Curriculum Compacting

Curriculum compacting is a differentiation strategy that was developed by Dr. Joseph Renzulli and Linda Smith in 1978. Designed to meet the needs of high-ability students, curriculum compacting streamlines schoolwork to a pace that corresponds with the student's ability in order to create a challenging learning environment. Curriculum compacting is a content acceleration strategy that enables students to skip parts of the curriculum they have already mastered and move on to more challenging content and activities (Tomlinson, 2003; VanTassel-Baska, 1997).

Differentiated Instructional Methods

“A teacher proactively plans varied approaches to what students need to learn, how they will learn it, and/or how they can express what they have learned in order to increase the likelihood that each student will learn as much as he or she can as efficiently as possible” (Tomlinson, 2003, p. 151).

English Language Learner

Students whose home language is not English (O'Neal & Ringler, 2010).

Enrichment

Curriculum as well as program delivery services that are more varied or has been modified or adapted to provide more depth or breadth for a student (Schiever & Maker, 1997).

Entrance Criteria/ Entrance Assessment/ Identification Measures

The tests or assessments given to determine eligibility or acceptance into a specific program (Assouline, 1997).

Ethnicity

Often defined as Hispanic or Latino and Not Hispanic or Latino. Race and ethnicity are social constructs rather than biologically based categories (Malley-Morrison & Hines, 2007).

Gatekeeper

A gatekeeper is a person who controls access to something; for example, via a city gate. In the late 20th century the term came into metaphorical use, referring to individuals who decide whether a given message will be distributed by a mass medium. Can also mean any way that access is restricted (Smith, 2001).

Genius

Extraordinary intellectual and creative power. A person of extraordinary intellect and talent: A person who has an exceptionally high intelligence quotient, typically above 140 (Morelock & Feldman, 1997).

Giftedness

“Gifted and talented students” means a child or youth who performs at or shows the potential for performing at a remarkably high level of accomplishment when compared to others of the same age, experience, or environment and who: (a) exhibits high performance capability in

an intellectual, creative, or artistic area; (b) possesses an unusual capacity for leadership; or (c) excels in a specific academic field (Texas Education Agency, 2009).

“Students, children, or youth who give evidence of high achievement capability in areas such as intellectual, creative, artistic, or leadership capacity, or in specific academic fields, and who need services and activities not ordinarily provided by the school in order to fully develop those capabilities” (U.S. Department of Education, 2002).

Homogeneous and Heterogeneous

Homogeneous is composed of parts or elements that are all of the same kind, essentially alike. Specifically in the field of gifted education – homogenous classes are those filled with only high ability students; heterogeneous classes have students of mixed abilities (Colangelo & Davis, 1997; Maker & Nielson, 1995).

Identification Measures for Gifted Programming

Any assessment used to identify a student for gifted programs. Range includes aptitude and achievement tests, teacher checklists, creativity surveys, anecdotal records, etc., depending on the purpose of the program (Assouline, 1997).

Innate Mental Capacity

Spearman’s *g* theory (shared by others) – cognitive ability is viewed as innate, stable and generalizable across all aspects of intelligent functioning (Nettelbeck & Wilson, 2005).

Intelligence/Aptitude

“It would be ridiculous to consider the score obtained from an administration of an intelligence test to be synonymous with intelligence. Nevertheless, a well-designed test can provide relevant and useful information about behaviors that typically are associated with ‘intelligence’” (Assouline, 1997, p. 91).

Spearman’s *g* theory – cognitive ability is viewed as innate, stable and generalizable across all aspects of intelligent functioning. No other form of measuring mental ability is as widely useful as *g* (Nettelbeck & Wilson, 2005).

Other theorists include motivation, creativity, and an extended list of potential expressions of intelligence (artistic expression, physical accomplishments, etc) (i.e. Gardner, 1983; Sternberg, 1985; Witty, 1962).

IQ

Intelligence quotient (Assouline, 1997).

See also Intelligence/Aptitude and Aptitude Tests.

LEA

Local education agency

Learning Disabilities (LD)

Neurological disorder that affects the brain's ability to receive, process, store, and respond to information. Learning disabilities (LD) are a group of disorders that can impact many areas of learning, including reading, writing, spelling, math, listening, and oral expression (National Center for Learning Disabilities, 2011).

Typically diagnosed, in part, by a significant difference between measured aptitude and achievement (Clark, 1997; Francis et al., 2005).

Mental Retardation/Intellectual Disability

A term for a person who has certain limitations in mental functioning and in skills such as communicating, taking care of him or herself, and social skills. These limitations will cause a child to learn and develop more slowly than a typical child. Children with mental retardation may take longer to learn to speak, walk, and take care of their personal needs such as dressing or eating. They are likely to have trouble learning in school. They will learn, but it will take them longer. There may be some things they cannot learn.

Aptitude test scores typically two standard deviations below the norm or IQ score of 70, based on an intelligence test with a mean of 100. Severe mental retardation measures as an IQ of 50 or below (Luckasson et al., 2002).

The term “mental retardation” has transitioned to other terminology such as cognitive disabilities or intellectual disabilities (Meservy, 2008).

Minorities

A smaller subset of a larger population, often with social or physical traits deemed less by the majority (Coughlin, 1997).

Motivation

The definition of motivation is to give reason, incentive, enthusiasm, or interest that causes a specific action or certain behavior (Rimm, 1997a).

Non verbal Intelligence

Intelligence that is manifested in the performance of tasks requiring little or no use of language (Lohman, 2005a).

Perfectionism

A propensity for being displeased with anything that is not perceived as perfect or does not meet extremely high standards (Clark, 1997).

Prodigies

A child prodigy is someone who at an early age masters one or more skills at an adult level before the age of 10. One criterion for classifying prodigies is: a prodigy is a child, typically younger than 15 years old, who is performing at the level of a highly trained adult in a very demanding field of endeavor (Feldman, 1993).

The term *Wunderkind* (from German: "wonder child") is sometimes used as a synonym for prodigy, particularly in media accounts, although this term is discouraged in scientific literature. *Wunderkind* also is used to recognize those who achieve success and acclaim "early" in their adult careers, such as Steven Spielberg, Steve Jobs, and Fred Goodwin (Clark, 1997; Colangelo & Davis, 1997).

Scholastic Aptitude

Cognitive abilities related to school achievement (Kärkkäinen, Rätty & Kasanen, 2009).

Segregated

Set apart or separated from others of the same kind or group (Colangelo & Davis, 1997; Maker & Nielson, 1995). Also referred to as tracking in some cases.

Social-Emotional Difficulties

Gifted children may develop asynchronously: Their minds are often ahead of their physical growth, and specific cognitive and emotional functions are often developed differently (or to differing extents) at different stages of development. One frequently cited example of asynchronicity in early cognitive development is Albert Einstein, who did not speak until the age of 2, but whose later fluency and accomplishments belied this initial delay (Silverman, 1997).

Success

Nine-week grades, as defined in this study. Students in the academically gifted English middle school program that earn less than an 80% for a given nine weeks are placed on

probation. The following nine-weeks their average must be above an 80% or they can be removed from the program.

Talent Search Programs

Intended to capture and identify the top margin of high-ability students, talent search programs exist throughout the United States. Most are based on the work developed by J. Stanley at Johns Hopkins University (Brody, 2005). Many talent search programs also offer rigorous academic programs, summer and distance education courses, and dual credit courses for their identified high ability students (Assouline & Lupkowski-Shoplik, 1997).

Teacher Subjectivity

Subjectivity means based on or influenced by personal feelings, tastes, or opinions. Characteristic of some teachers that sometimes do not call on, grade, treat or otherwise interact with all students equally – based on the influence of their feelings for particular students (Feldhusen, 1997).

Team

Part of the middle school concept, interdisciplinary teaming involving a core of teachers assigned to the same students. Thus, the team teachers (one English, reading, math, science, social studies) all teach the same group of students and typically have a common conference period to meet with parents and discuss strategies for working with individual students (Wilson, 2007).

Underachievement

Underachievement occurs when a child's performance is below what is expected based on the child's ability (Rimm, 1997a).

Vocational Education

Educational training that provides practical experience in a particular occupational field, as agriculture, home economics, or industry (Carr, 2009).

Organization of the Study

This study is organized into five chapters. Chapter One provides general background, the need for the study, the research questions, limitations, and definitions. Chapter Two presents a review of the literature focusing on an overview of intelligence theorists, aptitude and achievement tests, gifted education, special populations within gifted education, and the four tests used in the study. Chapter Three describes the methodological and data collection techniques that will be used in this study. Chapter Four describes the statistical results, and Chapter Five includes a discussion of those results with recommendations for further study.

Summary

This chapter introduced the background that forms the foundation of this study. Nearly 100 years of discourse among theorists and educators has created the field of gifted education (Jolly & Bruno, 2010). But the field can be a tangled web of semantics (Assouline, 1997), so this chapter provided definitions, clear background information, and identified theorists in the field.

The history of intelligence and gifted was explored, as well as the local history of the gifted programs. Nearly 100 years of data support that intelligence and giftedness are worthy of continued study. The programming at the participating LEA was most recently updated in 2006 and was evaluated for the efficacy of these updates.

Furthermore, the need for the study was delineated, the research questions and null hypotheses were stated, and the limitations and assumptions explained. A wide range of definitions were provided and the organization of the entire study was presented.

Beyond the foundation of definitions and history, this study intended to meet the needs, partially, of the fragile position of gifted education in our public schools. Sisk (1990) asserted, “the quality of the future leaders of our society is directly related to the quality of the educational experience provided for them” (p. 1). Accurately identifying students for gifted programs is a foundational piece that should be examined. This dissertation extended this issue one step further by providing recommendations about the different instruments used and which might be considered the strongest predictor of success.

Finally, Jolly (2004) noted the lack of published works exploring giftedness in theory and/or application, as well as its history. Dai and colleagues (2011) asserted that the field of gifted education must have more systemic research to move forward. It is the goal of this dissertation added to contribute to the field of gifted educational research.

Chapter Two

Review of the Literature

The review of the literature provides a foundation of background information upon which this study was constructed. The interest in evaluating the efficacy of assessments as tools to identify gifted students is present in the literature (Gustin & Corazza, 1994; Mattox, 1995; Niederer, Irwin, Irwin, & Reilly, 2003), though none shared the same exact measures or type of program.

No research was found that examined the relationship between the exact set of assessments used in this dissertation and student performance in a gifted English middle school classroom. The four entrance assessments evaluated in this study are the ones used by the participating school district. In an effort to provide a theoretical framework and basis for this research, the review of the literature begins by addressing the component and purpose of the study, then examining parallel studies. These parallel studies include research that is structured similarly but compares entrance criteria to achievement in different subject matter. The next sections construct a theoretical foundation of intelligence theorists, aptitude and achievement tests, and gifted education. These theories and research are the pillars upon which the current assessments were selected and the gifted program designed. Therefore, it is essential that they be explored. Finally, the last section delineates the difficulties in identifying gifted students within special populations. There has been a call to action within the gifted community in past years to accurately and fairly identify gifted learners, so the inclusion of this topic in the literature review is merited.

Components of Study

This review of the literature covered eight major topics. The first and second section identified the components and purpose of the study. The third section evaluated parallel studies in an effort to establish a research base in the absence of matching or more similar studies. These parallel publications vary in content or measures used but follow the same format of investigating the efficacy of qualifying assessments. The fourth division of this literature review focused on the theorists who studied intelligence. Since above-average intelligence is often included in the definition of giftedness (Assouline, 1997) and is the basis of the majority of gifted programs, a review of intelligence theorists was pertinent. The fifth section extended the work of the intelligence theorists to the tests created to measure intelligence, including the identification of gifted learners. All the tests discussed are commonly used and typically known among educators (Lohman, 2005b; Nasca, 1988). Most of them have been widely studied (Assouline; Lohman; Loyd, Forsyth, & Hoover, 1981; Gordon, Duff, Davidson, & Whitaker, 2010; Madaus & Rippey, 1966; Nasca; Newton, McIntosh, Dixon, Williams, & Youman, 2008; Robins & Jolly, 2011), used for many years and in a variety of settings, and researched at length. The sixth section established the history and current progress of gifted education. The seventh section addressed the specific needs of several special populations when identifying gifted learners. These special populations have at one time been considered underrepresented among gifted students and this reviewed some of that history as well as the current research findings. The final section is background information regarding the four entrance tests used in the local school district. The intent of this literature review was to provide an overview of related studies, intelligence theories, aptitude and achievement tests, gifted education, and underrepresented gifted student populations.

Purpose of Study

There is little agreement within the field of education regarding the construct of intelligence, with more than 100 different definitions (Johnsen, 1997). There is even less agreement about the definition of giftedness (Gottfredson, 2004; Konstantopoulos, Modi, & Hedges, 2002). Due to this lack of agreement, there is no universal identification or programming for gifted students (Louis et al., 2000). According to NAGC (2009), each state is allowed to determine how they interpret the federal definition of gifted, and most states even allow individual school districts to select their definition to determine their programming (NAGC, 2011).

Program and identification agreement.

It is critical to establish identification measures that match programming and theoretical purposes for our gifted learners (Louis et al., 2000; Lohman, 2005a; Nettelbeck & Wilson, 2005; Samuels, 2007; Sternberg & Clinkenbeard, 1995). Lohman asserted that admission to gifted programs should be “guided by evidence of aptitude for the particular types of advanced instruction that can be offered by schools” (p. 333). For example, Louis and colleagues expressed concern when students with above-average aptitude and high achievement are placed into highly creative or self-driven types of programs, such as Betts’ (1985) Autonomous Learner Model. If using Renzulli’s (1986) three-ring conception of giftedness, a district might determine if a child has strengths in the three areas: task commitment, creativity, and above-average ability. However, if the gifted program is exclusively an advanced academic course option with minimal creativity required, there is little match between the academic program and Renzulli’s identification model. According to Sternberg and Clinkenbeard, “in the absence of any model,

children are identified in a way that is one step beyond the toss of the dice” (p. 2). They believed the model should drive the decision-making regarding measures, instruction, and assessment (Hansen & Linden, 1990; Sternberg & Clinkenbeard).

Measurement of success.

This study examined the identification measures used in a large Texas school district (65,000 students). Gifted students in this district are identified using a variety of aptitude and achievement measures that match the academic programs offered. Every student is evaluated before entering the middle school program using four specific entrance assessments. Once enrolled in the program, students must maintain a minimum of an 80% nine-week average. Any student who has two consecutive grading periods with averages below an 80% may be dismissed from the program.

The teachers assigned to teach classes for gifted students have met the state minimum requirement of 30 hours of teacher in service training. Additionally, teachers attend a six-hour training meeting each fall and sometimes an additional six-hour training meeting in the spring. Topics addressed at those training meetings include differentiation within the gifted class, depth and breadth of the gifted curriculum, common struggles of gifted students, characteristics of a gifted classroom, and characteristics of an effective gifted teacher, among others. It is through these meetings that the district coordinator strives to ensure quality instruction in every gifted class in the district. Even if teachers are the most qualified, it was imperative that the program be assessed for its efficacy as a preliminary step toward investigating the relationship between the combined scores of the entrance assessments used in this district and student earned grades.

Evaluation of entrance criteria.

Gifted students have unique social and emotional needs (Cross & Coleman, 1993; Hollingworth, 1938). For example, many gifted students suffer from perfectionism, strained peer interactions due to their giftedness, and underachievement, to name a few (Colangelo, 1997; Cross & Coleman; Neumeister, Williams, & Cross, 2009). Identifying children as gifted and then effectively telling them that they “couldn’t cut it” with matching grades may be devastating (Colangelo). In an effort to accurately identify students for the gifted program, intervening factors (e.g., test anxiety, family trauma, and health issues) should be taken into consideration. However, it is equally important to avoid misplacing above-average students into gifted courses where they might not be able to perform well or adjust to the pace, depth or breadth of the curriculum.

Parallel Studies

Literature related to gifted assessment exists, but it is limited in many areas. No identical studies were found, but some parallel ones were discovered. When exact studies do not exist in the literature, parallel studies can form a solid foundation for research (Pyrzack & Bruce, 2000). These studies create a footprint that can be used to provide a relevant research pattern.

Since no research was found examining these exact entrance assessments for a middle school English gifted program, parallel studies were explored in the current investigation. These parallel studies examine the relationship between entrance criteria and achievement. The entrance criteria or subject matter may be different. In this literature review, both studies with

similar entrance criteria and different subject matter, and studies with different entrance criteria but similar subject matter were examined.

Spencer (1996) investigated the correlation between an achievement test and college grades. Newton et al. (2008) reported evidence of correlations found between shortened measures of intelligence to the Stanford-Binet Intelligence Test. Additionally, several studies have evaluated the predictive nature of both achievement and intelligence testing in a variety of settings and for a variety purposes (Benbow, 1992; Grossman & Johnson, 1983; Gustin & Corazza, 1994; Hunsaker, Abeel, & Callahan, 1991; Loyd, Forsyth, & Hoover, 1981; Nasca, 1988; Nettelbeck & Wilson, 2000; and Niederer et al., 2003). Further, several researchers have investigated the relationship between aptitude and achievement tests and success for English language learners (ELL) (Lohman, Korb, & Lakin, 2008), at-risk students (Bonner, 2000; Ford, 1995; Lohman, 2005b; Tyson, Darity, & Castellino, 2005), and females (Dai, 2002; Kerr, 1997; Nelson & Smith, 2001; Rebhorn & Miles, 1999; Reis, 2003; Roeper, 2003; Sadker, 2002; Schober, Reimann, & Wagner, 2004).

A number of studies have explored the relationship of various assessments with the students they identify (Frasier, 1997; Gustin & Corazza, 1994; Hunsaker, Abeel, & Callahan, 1991; Lohman, 2005a; Lohman, Korb, & Lakin, 2008). Some researchers have used correlation studies to determine the strength of the relationship between the two variables, while others have conducted regression studies (Grossman & Johnson, 1983; Mattox, 1995) to indicate the predictive validity of the instruments. Regardless of the statistical methodology used, numerous studies (Ford, 2004; Ford, Grantham, & Whiting, 2008) have investigated the connections

between the way in which students were identified for gifted programs and the intervening factors that are sometimes involved (e.g., gender, race, background).

While no exact studies were identified, several similar studies were found to be helpful. These mimicked the format and research methodology, specialized groups of students, or types of assessments used to determine the relationship with measured success. These parallel studies were helpful for finding patterns within the literature and field – for both further study and missing components.

Progressive Achievement Test (PAT).

The need to identify students for gifted programs extends beyond our borders. Niederer and colleagues (2003) examined the predictive validity of the Progressive Achievement Test (PAT) and a published problem-solving test to identify mathematically gifted children in New Zealand. The goal was to ascertain if primary school-age children could be identified using the PAT. Using the problem-solving test as entrance criteria, the researcher analyzed scores on the PAT to find correlations with the problem-solving test. Due to only 78% accuracy, using the PAT for this purpose was not recommended (Niederer et al.).

Scholastic Aptitude Test (SAT).

As one of the hallmark tests used to identify students who will succeed in college, the SAT has been evaluated for its predictive validity time and again (Bridgeman, Burton, & Pollack, 2008; Geiser & Studley, 2002; Mills & Barnett, 1992; Spencer, 1996; Stringer, 2008). The combination of the quantitative, verbal, and essay scores is used for college entrance to

numerous universities (Bridgeman et al.) and to determine National Merit Scholarship recipients using the Pre-SAT (Johnsen, 2006).

Further, the SAT has been used to identify and study extremely advanced gifted students, most notably by Julian Stanley (Brody, 2005; Stanley, 2005) through the Center for Talented Youth at Johns Hopkins University. Beginning with precocious 13- and 14-year-old university students in the 1970s, Stanley counseled them, help them choose their courses, and “provided considerable support for their emotional needs” (Brody, p. 1). There are now six major talent search centers in the United States (Lee, Matthews, & Olszewski-Kubilius, 2008). The use of the SAT-M to identify talented seventh grade students across the nation has been a foundation for talent search programs for decades.

Controversy.

Weighing in on this issue have been popular media outlets like Newsweek and ABC News (Bronson, 2009; Paulous, 2009). Both cited the long history of the SAT, its related validity testing, as well as the controversy surrounding its predictive nature. Controversy surrounds the test because some perceive test bias exists and others claim the sample size is too small (Bollag, 2005; Bradley, 2001; Gehring, 2002; Stringer, 2008). Others criticize the SAT for racial, socio-economic, and gender differences (Scherbaum & Goldstein, 2008; Stringer). Recent studies have been conducted with recent data sets. Item difficulty and item response theory, in association with the racial variables, were evaluated and have confirmed that some items function differently for the racial subgroups (Santelices & Wilson, 2010; Scherbaum & Goldstein).

While Santelices and Wilson (2010) criticized the SAT for its test bias, Stringer (2008) aimed his criticism at the acceptance of the SAT validity in several ways. One such concern is the lack of diversity among the studied population. Since “psychometric tests are generally designed to discriminate across a whole entry, and test-takers have been pre-selected using another criterion, a great deal of the variation in the test scores is lost, so comparisons may be across an unsuitably small range” (Stringer, p. 54). This means that the SAT scores are only varied as much as college applicants typically vary, which could result in meaningless data (Stringer). While no one argues that the entire range is not typically included in college entrance groups, it is wide enough to dispel this argument (Camara & Echternacht, 2000).

Despite claims regarding test bias by some, the SAT continues to be used to measure both verbal and quantitative reasoning to avoid lower predictive validity that comes with only nonverbal measures (Lohman et al., 2008). College academic success can be accurately predicted using the SAT (Camara & Echternacht, 2000). Camara and Echternacht suggested that, not only have there been hundreds of validity studies but that these studies “consistently find that high-school grades and SAT scores together are substantial and significant predictors of achievement in college” (p. 9). Bridgeman et al., (2008) concurred that numerous colleges and universities have been included in large-scale studies concluding that the SAT is valid. Interestingly, Geiser and Studley (2002) noted that students’ scores on the SAT II achievement tests are consistently better predictors, though slightly, of freshman grades than SAT I scores.

The use of the SAT to predict student earned grades in college is of great interest to universities, which need low attrition and high graduation rates. Furthermore, when combined with high school grades, the SAT has become the standard for assessing which students will be

able to compete in the college environment (Bridgeman et al., 2008; Geiser & Studley, 2002; Mills & Barnett, 1992; Spencer, 1996; Stringer, 2008). With a similar intent, this study sought to determine the strength of the relationship between the combined scores on the entrance criteria and the subsequent student earned grades in the gifted middle school English program.

Kindergarten Diagnostic Instrument (KDI).

Another parallel study involved the use of the Kindergarten Diagnostic Instrument (KDI) (Robinson & Miller, 1990). Mattox (1995) investigated the predictive validity of the KDI as a diagnostic instrument. Originally intended to reveal developmental delays, Mattox extended the KDI and used it to identify very young gifted students and to predict their later academic achievement. The number and nature of the research questions demanded a variety of statistical analysis – ranging from correlation to regression studies.

Mattox sought to use KDI scores to discriminate among students eligible for the Chapter I reading program (the name of the Title I program from 1981-1994) (Jennings, 2000) during first grade; second-grade achievement in math, reading, and total language as reflected on the California Achievement Test (*California*, 2009); and eligibility for the gifted math and the gifted language arts programs in third grade. Mattox found a positive relationship between the results on the KDI and second-grade achievement scores and the third-grade gifted language arts program, but a low negative association was found with the Chapter 1 reading program and the gifted math program.

Fraas and Frick (1992) studied the KDI to determine if children would eventually qualify for a first-grade reading intervention. Kurdek and Sinclair (2001) sought to examine the links

between age and gender differences and later achievement in reading and math. In addition to the original intention of the KDI, these various studies all demonstrate its efficacy to identify students with future high achievement.

Math and science.

The final segment of parallel studies includes studies that are using assessments to determine achievement in math and science rather than an English gifted program. Gustin and Corazza (1994) explored the predictive validity of the SAT-Mathematics (SAT-M) and SAT-Verbal (SAT-V) to predict success in science achievement. The SAT was chosen in this talent search program because it would provide a sufficient upper limit to distinguish between students scoring at or above the 97th percentile (Gustin & Corazza). Gustin and Corazza documented the success of the SAT used as a primary source for identifying talented youngsters. Using a Pearson Product-Moment Correlation, a relatively high correlation was found with the SAT-M and even higher with the SAT-V with science achievement scores (Gustin & Corazza). Additionally, Spencer (1996) noted that the SAT-M test scores do predict college chemistry grades rather accurately. These are powerful studies that verify that the talent search programs continue to necessitate the use of the SAT as a measurement tool to identify gifted youngsters. Additionally, the correlation between achievement scores and student earned grades, even in a different content area, was very valuable as a parallel study.

Benbow (1992) and Lupkowski-Shoplik and Assouline (1994) extended the use of the SAT further. Benbow argued that in addition to its use as an identifier of precocious adolescents, the SAT could be used to assess variances among individuals regarding development and to predict exceptional academic achievements. To this end, the SAT, used with seventh and eighth

graders, can discriminate among those attaining scores in the top 1%. Benbow found that among the students who scored within the top 1%, the SAT accurately predicted the level of success those students would have in the following ten years. “This indicates that there is usefulness in having high-ability students take the SAT-math at an early age, as more than 100,000 students now do annually (Benbow, p. 59).

Lupkowski-Shoplik and Assouline (1994) examined individuals with mathematical precocity using qualitative means to determine programmatic options for elementary-level gifted children. They discovered that oversights are made when program directors did not consider all the factors like gender, advanced programs, and an accelerated rate, to name a few.

While these studies are instrumental in the use of measures already in place to identify students with abilities far above average, it is specifically geared toward finding achievement in the areas of mathematics and science, not English. Furthermore, Benbow (1992) was using the SAT-M to identify gifted youngsters in the seventh and eighth grades; this dissertation study evaluated younger students for an English gifted course.

Parallel studies build a foundation upon which further studies can be built. The SAT measures students’ aptitude for college performance by measuring mathematics, reading and writing skills and is found to correlate highly with college freshman GPAs (Bridgeman et al., 2008). Other studies have compared student achievement in a specific content area with aptitude and achievement scores. These parallel studies are relevant even though they might not match exactly to the current investigation.

Intelligence Theorists

Few topics are more controversial in American public schools than the concept of intelligence (Gottfredson, 2004; Neisser et al., 1996; Weinberg, 1989). In the past three decades, there have been an increasing number of conversations about the type of intelligence and the different forms of accomplishment (Gottfredson). These can be grouped in a variety of ways, but for the purposes of this literature review they have been categorized into three groups.

The first group is comprised of those theorists who use a psychometric approach like Spearman, Guilford, and Thurstone – which are further divided into a second group, “splitters” or “lumpers” according to Mayr (1982). The third group includes theorists who ascribe to a multiple intelligence viewpoint like Gardner and Sternberg.

Psychometric approach.

Mayr (1982) divided intellectual theorists into the two divisions of “lumpers” and “splitters”, noting that some hold intermediate positions (e.g., Horn, 1982; Vernon, 1940) as neither lumpers nor splitters. “Lumpers define intelligence as a general, unified capacity for acquiring knowledge, reasoning, and solving problems that is demonstrated in different ways...” (Weinberg, 1989, p. 98). Most notable in this category is Charles Spearman (Norton, 2006; Williams, Zimmerman, Zumbo, & Ross, 2003) who built upon the work of Binet, Simon, and Terman. Spearman’s *g* – the general intelligence factor was a major development of his time and was instrumental in attempts to explain what intelligence tests really measure (Norton; Weinberg). Spearman’s *g* has been highly criticized, but also lauded as a strong predictor for many facets of individual accomplishment and learning (Kane & Brand, 2006).

The other group of psychometric theorists consists of the splitters. This group includes theorists like J. P. Guilford (1982) and Louis Thurstone (1946), and uphold that intelligence is composed of many separate mental abilities that operate more or less independently from one another (Weinberg, 1989). Guilford argued that there are more than 150 factors of intelligence. “Carroll (1993) provided the most detailed description yet available of the psychometric structure of human intelligence...identifying a complex model that encompasses a strong general ability, together with eight or nine additional broad forms of different intelligences” (Nettelbeck & Wilson, 2005, pp. 615-616). The combined research of Carroll, Cattell, Horn and Noll, and McGrew impacted the development of such measures as the Stanford-Binet and Wechsler Intelligence Scale for Children (Nettelbeck & Wilson). Thurstone concluded that there are special abilities that contain some central energizing factor, which promotes the activity of all these special abilities. Additionally, Thurstone expressed concerns over Spearman’s *g*, noting that not all expressions of intelligence could be explained by the general intelligence factor. Other splitters regard the human intellect as a complex set of natural abilities, not described or measured as easily as Spearman, Terman, or Binet would have argued (Weinberg). Most recently, Dweck (2010) has been shifting the construct of intelligence and giftedness away from a fixed mindset, an entity one is born with; to a growth mindset, whereby intelligence develops over time and is incremental in its nature.

Multiple forms of intelligence.

Howard Gardner is another splitter theorist, but approaches intelligence quite differently than most. He asserted that intelligence is evidenced in multiple independent domains. Currently Gardner has identified nine published intelligences, but “the list of current intelligences remains

a work in progress,” and Christodoulou (2009) projected more may be recognized in the future. Unlike other splitters who recognize intelligence in various domains, Gardner’s theory was not as firmly steeped in the “history of theorizing about the hierarchical structures of intelligence” (Weinberg, 1989, p. 99). Indeed, Gardner has been criticized for his lack of empirical data to support his theory (Sternberg, 2008).

Robert Sternberg’s Triarchic Theory of Intelligence (1985) places him among the splitter theorists, but vastly different than Gardner. Sternberg (2008) has criticized Spearman’s *g* because it ranks students on a one-dimensional scale, and he has criticized Gardner because there is an inadequate amount of empirical data supporting his multiple intelligence theories. Sternberg’s theory is far more universal (both in processing and mechanics) and based on three cornerstones (Weinberg, 1989). Sternberg’s theory is intended to assist with teaching successful intelligence so students can compensate for weaknesses and capitalize strengths. The three components of his theory are: (a) Intelligence must be interpreted within the socio-cultural context since intelligence can be recognized differently in various cultures; (b) Intelligence can be synthesized into two tasks: the ability to learn from new as well as previous experiences, developing expertise and automaticity; and (c) Intelligence requires information-processing skills and strategies (Weinberg). These three components are harmoniously united to balance analytic, creative and practical abilities (Sternberg).

Psychometric tests are typically used to evaluate aptitude/achievement and can inform professionals about a child’s capabilities (Nettelbeck & Wilson, 2005). Others have developed alternate assessment tools (e.g., Sternberg’s Multidimensional Abilities Test), and these may be used to supplement or substitute for traditional IQ measures. Despite the ongoing controversy,

which will likely remain (Weinberg, 1989), the field has been dominated by psychometric tests because they are useful in educational settings for a variety of student needs (Nettelbeck & Wilson), and because they can be “valid and useful gauges of student potential” (Gottfredson, 2004, p. 36).

Aptitude and Achievement Tests

The assessment of children’s aptitude is a foundation piece in most schools. Measuring intelligence is merely semantics (Assouline, 1997). “It would be ridiculous to consider the score obtained from an administration of an intelligence test to be synonymous with intelligence” (Assouline, p. 92). Most theorists recognize that all cognitive tests measure developed abilities in some way (Lohman, 2006). Lohman explained that “most novices believe that ability is innate and achievement acquired, whereas experts see the two as different aspects of the same thing” (p. 32).

Many theorists recognize that semantics are often the only thing truly separating these two constructs as they are typically assessed in our educational system (Lohman, 2006). However, because students do display a discrepancy between these two measures, as previously used to identify special education students, this study treated them as interrelated, but independent measures.

Identification of gifted learners.

Historically, many gifted programs focused on aptitude and achievement test scores as measures of giftedness; consequently, they are typically used as entrance criteria for many gifted programs. According to the NAGC website, in 2009 there were approximately three million

academically gifted and talented students in U.S. classrooms - a number that demands identification and programming be evaluated stringently because they are important elements in creating a quality education for gifted students (McGonagill, 1997; Sisk, 1990; Sternberg & Clinkenbeard, 1995). “But the complex nature of identifying gifted students remains a challenge for many states and districts” (Samuels, 2007, p. 2). Samuels asserted that this complexity means that there is “no silver bullet on identification” (p. 3). Dweck (2010) argued that giftedness is not static, but instead can be developed over time based on opportunities. Further, Cross (2009b) recognized that, in addition to innate intelligence, the development and life experiences of the gifted child impact his or her potential achievement and accomplishment.

Concerns about aptitude and achievement testing have been present almost as long as these type of tests have been used (Benbow, 1992; Gottfredson, 2004; Nettelbeck & Wilson, 2005). Concerns range from test bias against females and minorities, to use of tests to perpetuate social and economic injustice and misuse of test information (Benbow). Regardless of any controversy, “we continue to advocate the use of IQ and similar aptitude/achievement tests because, given current knowledge limitations, they are the best tools available for predicting important future educational and other significant life outcomes” (Nettelbeck & Wilson, p. 625). These predictions can be made using a variety of instruments and are typically from one of two approaches.

Psychometric approach.

Grossman and Johnson (1983) evaluated the validity of the Otis-Lennon (Otis & Lennon, 1997) and Slosson (Erford, Vitali, & Slosson, 1999) to predict achievement on the Stanford Achievement Test. Using multiple regressions, the findings indicated high levels of correlation

among these tests and final recommendations were to rely on the commonly used Otis-Lennon as a “valid indicator of academic performance and thus of cognitive ability for a population of students who are being considered for inclusion in a gifted program in the public schools” (Grossman & Johnson, p. 621). Other studies have been conducted to verify the predictive validity of several popular aptitude and achievement tests.

In an effort to quickly and easily identify students for gifted programming in the public school arena, shortened individualized aptitude tests have been used. Newton et al. (2008) investigated the efficacy of the Woodcock-Johnson Tests of Cognitive Ability, Third Edition Brief Intellectual Ability score (Woodcock, McGrew, & Mather, 2001); the Stanford-Binet Intelligence Scale, Fifth Edition Abbreviated IQ (Roid, 2003); and the Kaufman Brief Intelligence Test IQ Composite (Kaufman & Kaufman, 2004) in predicting giftedness as compared to the Stanford-Binet Full Scale IQ score (Roid).

The abbreviated tests are shorter, more easily given to small groups, and commonly used to identify qualifying gifted students. A Pearson correlation study revealed positive and significant relationships among the four scores (Newton et al., 2008). The Woodcock-Johnson was most closely aligned with the Stanford-Binet Full Scale scores, while the Kaufman was least accurate in identifying giftedness (Newton et al.).

Newton et al. (2008) included a valuable perspective regarding the importance of accurate identification measures for programs following the updated Marland Report in 1993, “which stated that gifted children would be provided with educational opportunities that facilitate and enhance academic and intellectual progress” (p. 535). Whether a full-scale or brief intelligence test, the instruments used to determine educational decisions should be evaluated

strenuously (Johnsen, 1988). Without valid measures to screen and assess students for our gifted programs, educators are left to make uninformed decisions (Newton et al.).

Loyd and colleagues (1981) explored the relationship between achievement test scores and academic achievement in high school and college. Using the Iowa Test of Basic Skills and the Iowa Test of Educational Development (Hoover et al., 2001), they studied the predictive validity of such measures on college students' academic performance. The study "supports the contention that proficiency in basic skills at the elementary level as well as the high school level ... does have a significant relationship to high school and college academic success" (Loyd et al., p. 1124).

Another study of interest examined the Secondary School Admissions Test (SSAT) to "institute a talent search for elementary school-age children" (Mills & Barnett, 1992, p. 155). The SSAT measures verbal and mathematical ability, as well as reading comprehension. After identifying the specific test form to use and an appropriate cut-off score, Mills and Barnett found using the SSAT was accurate in "identifying elementary age students who have exceptional verbal and quantitative reasoning ability" (p. 155).

Multiple forms of intelligence.

Howard Gardner asserted that intelligence is evidenced in multiple independent domains. Typically, students are assessed by use of a checklist or survey to determine areas of intelligence (Weinberg, 1989). Gardner's theory is not as firmly steeped in the "history of theorizing about the hierarchical structures of intelligence" (Weinberg, p. 99), and his theory has not been

validated (Gottfredson, 2004; Nettelbeck & Wilson, 2005). The lack of empirical evidence to support Gardner's theories has brought a great deal of criticism to his work (Sternberg, 2008).

Sternberg's Triarchic Theory of Intelligence (1985) is vastly different than Gardner's. Sternberg's theory is based on three cornerstones: (a) Intelligence must be interpreted within the sociocultural context since intelligence can be recognized differently in various cultures; (b) Intelligence can be synthesized into two tasks: the ability to learn from new as well as previous experiences, developing expertise and automaticity; and (c) Intelligence requires information-processing skills and strategies (Sternberg, 1997; Weinberg, 1989). These three components should harmoniously unite to balance analytic, creative and practical abilities (Sternberg, 2008).

The field of theories of intelligence and its measurement is relatively young – slightly more than one hundred years old – and yet, it has made momentous strides. From the beginnings, with Galton, Binet, Simon, and Terman, the empirical data and theories regarding intelligence have grown exponentially. For example, there are now more than one hundred different definitions of intelligence with nearly as many instruments to measure it (Johnsen, 1997). “Little consensus exists among educators regarding how to define the term intelligent and gifted and how to do so equitably” (Ford & Grantham, 2003, p. 218). Even Gardner and Sternberg concurred that the *g* factor exists, only disputing “its preeminence among mental abilities” (Gottfredson, 2004, p. 37). These two theorists might wish for distinct forms of mental abilities to equally share the spotlight, but a century of research has revealed that the concept of *g* is the most widely useful (Gottfredson).

Schools have been using some form of aptitude or achievement test for decades, applying the information gained to assist in identifying specific needs of students. For example, these

types of tests are used to identify deficits so that struggling learners can be guided toward success, and some researchers have found those same measures can rather accurately predict students who will process and retain information more quickly than most – the gifted (Baker, 2001; Nettelbeck & Wilson, 2005). Consequently, American schools have devised a system of using aptitude and achievement tests to identify and qualify students for specialized programs. The programs for gifted students intend to provide enrichment, acceleration, or compacted curriculum commensurate with the pace most gifted students can handle. When used appropriately, the aptitude and achievement tests provide educators with a wealth of information pertinent to school success for many children.

Controversy

Aptitude and achievement tests have been criticized for years, but “psychologists continue to see them as valid and useful gauges of student potential” (Gottfredson, 2004, p. 36). One concern stems from those individuals who might misinterpret the scores to be representative of a person’s entire range of skills and abilities (Weinberg, 1989). Some argue that these tests lack “prescriptive utility” (Nettelbeck & Wilson, 2005, p. 609), and others are concerned with the misuse of IQ testing (e.g., inappropriate testing situations), as occurred in the early part of the 20th century (Nettelbeck & Wilson; Thorndike, 1975). For instance, Terman’s original testing was used to determine vocation, management-training program participants, adoption viability, suitability for military leadership, as well as to make generalizations about racial differences (Jolly, 2008; Thorndike; Weinberg). Critics have expressed concern that the original intention to exclude some students from public education negates the current uses to identify students with exceptionalities (Weinberg). Additionally, some critics have argued that testing programs have

negative effects on students, especially when used inappropriately or interpreted inaccurately (Hansen & Linden, 1990; Weinberg).

Probably the most notable condemnation of these tests falls to the racial, socio-economic, and gender inequities that Herrnstein and Murray's book *The Bell Curve* (1994) brought to light (Pyryt, 1996). However, the response to *The Bell Curve* was notable, as many theorists in gifted education argued for the validity and reliability of IQ tests (Pyryt; Sternberg, 1996).

“We continue to advocate the use of IQ and similar aptitude/achievement tests because, given current knowledge limitations, they are the best tools available for predicting important future educational and other significant life outcomes” (Nettelbeck & Wilson, 2005, p. 626). But most importantly, in response to the elitism of *The Bell Curve*, Sternberg (1996) reminded us all that “however it is defined, intelligence is only one attribute of human beings and one attribute leading to certain kinds of success ... they are not measures of human worth” (p. 15).

Additionally, there has been some controversy (Lohman, 2005b; Lohman et al., 2008) regarding the use of non-verbal aptitude tests to identify gifted students. School success depends on students' ability to use language to express thoughts and understand others, so students who can succeed in formal schooling are those who are best able to reason verbally (Lohman, p. 344). A primary concern highlighted by Lohman was that “nonverbal reasoning tests do not identify the same students as tests of verbal or quantitative reasoning abilities” (p. 347). Lohman divided potential gifted students into two prospective groups: high-accomplishment, which were students who currently demonstrate “academic excellence in a particular domain”; and high-potential, which were those likely to be successful if they are willing to work hard for it. Since these two groups need different instructional programs, Lohman suggested that mixing them is

counterproductive and that they should not share the same label. Furthermore, Lohman indicated that academically talented minority students would be best identified using these two groups, since minority students tend to have difficulty qualifying, for whatever reason, for traditional gifted programs that only assess high-accomplishment students.

Gifted Education

Nationally.

The federal government began to take an interest in giftedness during the 1840s and again in 1885 when Sir Francis Galton attempted to differentiate between the mental abilities of various individuals (Thurstone, 1946). However, a historical perspective of gifted education cannot be discussed without evaluating Lewis Terman's work (Delisle, 1999; Jolly, 2008; Nettelbeck & Wilson, 2005). Terman was a "major catalyst" in the development of our collective perception of gifted individuals (Sisk, 1990, p. 2) because of the volume of his work (e.g., 1920, 1921, 1925, 1939) (Konstantopoulos et al., 2002; Sisk; Thorndike, 1975; Weinberg, 1989). Terman identified over 1,500 gifted children and conducted a longitudinal study that extended into adulthood (Jolly, 2004; Sisk). Beginning in the 1920's, Terman reversed public opinion about highly intelligent individuals; his study found these gifted individuals strong, healthy, socially active and productive members of society (Sisk).

Lohman (2006) recently continued some of Terman's work on perceptions and awareness of high-ability individuals with similar results. Additional contributions from Terman have laid a foundation for researching, assessing, educating and servicing gifted individuals (Konstantopoulos et al., 2002). Among many findings, Terman discovered that instructional

acceleration is beneficial at all levels (Sisk, 1990). Following Terman's death in 1989, Stanford University appointed Professors Robert and Pauline Sears to continue the study of these gifted individuals (Leslie, 2000; Sisk). Following the Sears' deaths, former Stanford provost and vice president, Al Hastorf, was appointed the third director of the Terman study (Leslie). Hastorf asserted that the longitudinal study of Terman's originally identified gifted individuals "will continue until the last one dies" (Leslie, "The Genetic Study," para. 7). Hastorf's successor has not been appointed following his death in September 2011.

Interest in giftedness peaked again during World War II (McClellan, 1985). The National Science and Foundation Act of 1950 (P.L. 507) was the first time when money was specifically funneled for programs targeted at gifted students (McClellan). Following Sputnik's launch in 1957, the United States witnessed a surge in science and mathematics education reform (Jolly, 2009; Roberts, 1999; Stewart, 1999). By 1969, when the United States sent astronauts to the moon, Congress mandated a study to evaluate the current needs of gifted and talented students in the United States (Jolly; McClellan).

Consequently, formal gifted education became far more widespread in the United States following the publication of the Marland Report (Marland, 1972). Gifted programs began appearing in public schools throughout the nation (Jolly, 2009; McClellan, 1985). At the same time, theorists (e.g., Renzulli, Spearman, and Sternberg) were grappling with a singular definition of intelligence and giftedness. Lack of agreement led to a wide variety of programs across the nation, in part because the federal definition does not mandate how each state implements gifted education (NAGC, 2011; Samuels, 2007).

During this same period of time, Julian Stanley began his work with gifted students at Johns Hopkins University (Brody, 2005; Stanley, 2005). Stanley began by following and assisting precocious students who were admitted to the university at very young ages. This led to him to conduct the second longitudinal of gifted youngsters in the United States, following Terman (Stanley). The study of these, and subsequent gifted children, prompted many of the talent search programs, and continues to this day (Brody; Assouline & Lupkowski-Shoplak, 1997; Stanley).

As the 1980s began, theorists and state gifted program directors debated what the appropriate requirements, as well as what the best programming, might be for identifying gifted students. The National Commission on Excellence in Education report, *A Nation at Risk* (1983), reported half of gifted student's school achievement did not match their tested ability (Delisle, 1999; Sisk, 1990). These concerns are still prevalent today, and controversy continues regarding the identification of gifted students. "Little consensus exists among educators regarding how to define the term intelligent and gifted and how to do so equitably" (Ford & Grantham, 2003, p. 218).

During the 1980s, there were several legislative acts (e.g., Omnibus Reconciliation Act of 1981, Regulations for the Educational Security Act of 1984) defining areas that grants could be given within the field of gifted education. These reflected an extension of the Marland definition by including specific fields of interest and incorporating measures to include students that may come from historically underrepresented areas at that time (McClellan, 1985).

Perhaps the most moving change occurred following the Jacob K. Javits Gifted and Talented Students Act of 1988, "which marked the culmination of the efforts of many dedicated

gifted education proponents” (Ford, 1995, p. 52). This piece of legislation was created to turn attention and financial support to the issues of gifted racial minorities, economically disadvantaged, English language learners, and students with disabilities.

While the heightened awareness of the underrepresented groups has assuaged the problem of identifying gifted students from these groups in some ways, current literature indicates it is still a concern across the country (Bonner, 2000; Ford, Grantham, & Whiting, 2008; Raborn, 2002; Robbins, 1991; Tyson et al., 2005). “Few school districts in the United States have successfully recruited and retained culturally diverse students in programs for gifted students” (Ford & Grantham, 2003, p. 217). Further challenges have faced gifted education most recently with the passage of No Child Left Behind Act in 2001 (NCLB)(USDOE, 2002). “Ironically, a piece of legislation meant to serve all students, instead focuses on proficiency and accountability, which leaves little room for the needs of the gifted” (Jolly, 2009, p. 49).

Even with additional information about the identification of minority students, gifted program directors are left to make decisions about definitions and programming for their district, because individual school districts typically function separately from one another. (NAGC, 2009) “Identification of gifted children ideally is tied to the theory of giftedness to which one subscribes. Pragmatically, however, it more often is tied to the particular program for which one is seeking participants” (Louis et al., 2000, p. 295). Therefore, the range of gifted programs across our nation is still very diverse (NAGC, 2010a).

State and local education agency (LEA).

Texas began implementing gifted programs in the 1970s in response to the Marland Report. By 1995, Texas was identified as one of three states with well developed policies for the identification of gifted students (Coleman & Galagher, 1995). Coleman and Galagher found Texas to have longevity of leadership and policies that led to special services and programs for gifted learners. These researchers pointed out that litigation led to a more demographically diverse gifted population in Texas than other states by the mid-1990s. Baker (2001) asserted that the combination of the Jacob K. Javits Act in 1988 and the National Excellence report in 1993 (USDOE) invigorated the field of gifted education. Texas is noted as a leader among the state gifted programs in both policy and programming for gifted students (Baker).

The local district began identifying third-grade students in the 1977-78 school year and added a year of gifted programming as those students aged, according to the current program director. The first year that sixth-grade middle school gifted students were grouped homogenously in English classes – rather than being grouped with all high achievers – was in 1984. According to the current district program director, the assessments used to determine qualification for the program may have changed over the years, but the four tests used to admit middle school students to the English gifted program and the criteria for remaining in the program has been in place at least since 1995. The most recent change occurred in 2007 when elementary gifted students were no longer promoted to the middle school gifted programs without retesting.

Special Populations

Empirical gifted studies have expressed ongoing concern about the numbers of students who are excluded from gifted programs due to intervening factors (Hunsaker et al., 1991). This concern, as well as the history of these special populations and their inclusion in gifted education programs, is addressed in this literature review. Gifted education has been evolving and the state of special populations can drive the research, entrance assessments, and programming (Johnson, Karnes, & Carr, 1997).

Minority students (specifically African American students) have been the focus of numerous studies for many years, but joining the cadre of publications are ones focused on Native Americans, Hispanics, English language learners, and students identified as at-risk. Additionally, girls have been the focus of research for many years. Roeper (2003) and Reis (2003) have both contributed to the foundation of work studying the participation of girls in gifted programs, especially those in typically male-dominated fields like math and science. Within each of these groups there also exists a population of special education students who are the most rarely identified for inclusion in gifted programs (Yssel, Prater, & Smith, 2010).

Hunsaker and colleagues expressed concern in 1991 when collecting data from school districts across the nation regarding the attempts (or lack thereof) made to identify gifted learners from specific populations such as limited English speakers, low socioeconomic-status students, or students with disabilities. These researchers found that most districts only moderately addressed those in the racial/ethnic category. The United States Department of Education (USDOE) produced the *National Excellence: A Case for Developing America's Talent* in 1993 to explain the current status of gifted and talented education and found similar results, as talented

children in several categories were overlooked. “These included culturally different children (including minority and economically disadvantaged students), females (who are underserved in mathematics and science programs), students with disabilities, high potential students who underachieve in school, and students with artistic talent” (USDOE, p.1). In each of these groups, we have an emergence of students with disabilities (e.g., attention deficit hyperactivity disorder, visual impairments, attention deficit disorder) – many who could be deemed as “twice exceptional” (Coleman, Harradine, & King, 2005). These students might have an emotional disability or a learning disability in reading, but be gifted in mathematics. Students with disabilities struggle to be identified as gifted and find success within gifted programs (Yssel et al., 2010).

Even though nearly two decades have passed since these reports were published, things have not changed very much. This was further echoed in the 2008-2009 State of the Nation report from NAGC (2009), “In the 40 years since the lunar landing ignited so many young minds, our commitment to academic excellence has waned, leaving gaps in our ability to identify and appropriately challenge high-ability students from all backgrounds” (p. 4). This underscores that the problems remain in special groups, “we have holes in a piecemeal collection and resources that result in a loss, not only to the students, but also to the nation” (NAGC, p. 4). Several researchers have investigated specific groups in an effort to ensure that identification measures are fair, unbiased, and that specific measures did not exclude individuals of any group (Frasier, 1997; Lohman, 2005a; Lohman et al., 2008). Milgram (2000) asserted that there were limited cross-cultural research studies completed in the gifted arena. In addition, schools face the challenge of educating students who are not originally from America – students who need language support in the classroom.

English language learners (ELL).

Lohman and colleagues published interesting findings in 2008 regarding the validity of commonly used assessment tools for ELL students. Their study compared the validity of three nonverbal tests used to identify academically gifted students. Since many program directors select nonverbal reasoning tests “to screen all students for inclusion in programs for the gifted because it is thought that such tests level the playing field for ELL and non-ELL children” (Lohman et al., p. 276), it is imperative that the measures be well-studied and that they demonstrate an accurate prediction of academic giftedness.

The Raven Standard Progressive Matrices (Raven), the Naglieri Nonverbal Ability Test (NNAT), and the Cognitive Abilities Test (CogAT) were measured with a population of ELL students. Lohman et al. (2008) warned that program directors not view nonverbal tests as a perfect measure of innate ability blind to culture, education, or experience in an effort to identify academically gifted children. These nonverbal tests can provide valuable information and help identify gifted youngsters, especially when combined with measures of quantitative reasoning and spatial ability (Lohman et al.). A comprehensive identification plan can alleviate the problems of nonverbal tests as long as a broad range of abilities is evaluated (Lohman et al.). The entrance criteria for this gifted English middle school program reflect that finding; the NNAT is a non-verbal test used in conjunction with additional measures to evaluate a broad range of abilities needed for success in this program.

According to Lohman (2005a), “predictors of achievement in reading, mathematics, social studies, and science are the same for White, Black, Hispanic, and Asian American students” (p. 344). Lohman suggested that test norms be compared to those of students of the

same background, especially students who are older or younger than grade peers. However, Lohman (2005b) warned that “selecting students for gifted and talented programs on the basis of nonverbal reasoning ability would admit many students who are unprepared for ... advanced instruction in literacy, language arts ... or other content-rich domains” (p. 116). The need to identify underrepresented groups stretches beyond the use of nonverbal testing.

Minorities.

In contrast to the dearth of empirical literature on giftedness, research about minorities and the achievement gap abounds. And disparity of minority students in gifted programs remains “despite increased knowledge about the different effects that cultural backgrounds, environments, and socioeconomic status may have on the expression of gifted behaviors” (Frasier, 1997, p. 498). Studies have found that “high-achievers will often downplay or even mask their talents and accomplishments” if the cultural norm does not support achievement or academic success (Markus, Steele, & Steele, 2000; Robbins, 1991, p. 1). Further complications arise when parents are apprehensive about including their child in a gifted program that is not culturally diverse (Ford, 2004) and the community exerts peer pressure about “acting White” (Ford et al., 2008).

Mills and Tissot (1995) suggested using the Raven Advanced Progressive Matrices test as a nonverbal evaluation instrument to identify academically gifted students from minority populations. Burris and Welner (2005) provided valuable conclusions about increasing achievement among minority students by eliminating tracking, but there was no application of that information to gifted identification specifically. Another improvement has been the increasing use of multiple criteria to identify students (Frasier, 1997). Others have used

enrichment after-school programs to build opportunities for students from minority populations (Native American) to “try out” gifted material and consequently develop positive attitudes about intelligence and gifted programs leading to an increase in minority qualification and participation (Raborn, 2002).

Ford and Grantham (2003) insisted that the underrepresentation is due to teachers who view minority students through a lens of deficit thinking – these educators “hold negative, stereotypical and counterproductive views about culturally diverse students and lower their expectations of these students accordingly” resulting in lower achievement across the board (p. 217). Ford and colleagues (2008) agreed that culturally and linguistically diverse students fall into the category regularly viewed through deficit thinking, noting that this hinders underrepresented students from qualifying and maintaining success in gifted programs. Since nearly all school districts use test scores for placement decisions for gifted programs (Colangelo & Davis, 1997), the lowered achievement of the minority due to teachers’ perceptions and expectations is to blame (Ford & Grantham). Ford and Grantham are quick to add that there are other contributing factors like the lack of agreement of a definition of intelligence, lack of agreement of a definition of giftedness, policies and practices used in gifted education, teacher multicultural education, teacher assessment education, communication with diverse communities, and student perceptions about gifted education.

Bianco’s (2005) research revealed that teachers exposed to a variety of student data routinely do not recommend students for gifted programs if a special education label was included – even when all other data were identical. Increasing numbers of children are being identified as both gifted and disabled – ranging from emotional and learning disabilities to forms

of autism (most notably Asperger's Disorder) (Amend, Schuler, Beaver-Gavin, & Beights, 2009). There are also students whose primary exceptionalities include other challenges such as physical disabilities, visual and hearing impairment or even attention deficit hyperactivity disorder. These twice-exceptional students are often inappropriately excluded from gifted programs (Coleman et al., 2005).

The school district in this study has faced and overcome many of the challenges regarding identifying underrepresented gifted students. As a result, the program director continues to see a rise in the identification and placement of twice exceptional students in the local program.

At risk students.

The at-risk indicator codes from the Texas Education Agency (2010) give 13 different criteria for identifying a student as being at-risk for dropping out of school. The 13 indicators are for any student who:

1. Is currently in one of the primary grades (PreK through third grade), but did not perform satisfactorily on a readiness or assessment instrument that during the year;
2. Earned below a 70 average in two or more core subject areas during the current or previous year;
3. Was retained in any grade level;
4. Did not perform satisfactorily on the state assessment during the current or previous year;
5. Is pregnant or is a parent;
6. Has been placed in an alternative education program during the current or previous year;

7. Has been expelled during the current or previous year;
8. Is currently on parole, probation, deferred prosecution, or other conditional release;
9. Was reported to have previously dropped out of school;
10. Is a student of limited English proficiency;
11. Is in the custody of the Department of Protective and Regulatory Services or has been referred to an officer of the juvenile court or law enforcement official;
12. Is homeless;
13. Resided in a detention facility, substance abuse treatment facility, emergency shelter, psychiatric hospital, halfway house, or foster group home in the current or previous year.

Additionally, students are considered at risk locally if they have a low socio-economic status (SES) routinely defined by qualification for the free or reduced-priced meal program in the schools (Ford et al., 2007). This typically places a student in an at-risk database used by the schools to identify and assist these students as needed. This is another common factor affecting student performance (Matthews, 2006).

Students at risk for dropping out constitute another underrepresented group (Matthews, 2006). There has been some debate in gifted education regarding the percentage of gifted students who drop out of school. The Marland Report in 1972 claimed that 18% of drop-outs are gifted students. Other studies report gifted students comprise nearly a third of all dropouts (Matthews; Renzulli & Park, 2000). Matthews pointed out that research in this area is difficult as the operational definitions of both giftedness and drop out are not clearly defined, causing the variance to be great. When studied, gifted students drop out for some of the same reasons as their peers: pregnancy, jobs that interfere with school, underachievement, and so on. (Hansen & Toso,

2007). However gifted students also cited frustrations with busywork, teachers who were disrespectful, and a generalized disconnect from school, their peers, and teachers as reasons for dropping out (Hansen & Toso; Renzulli & Park).

Regardless of the actual percentage of students leaving the school setting, only a few leave due to early college entrance or other acceptable reasons; many simply leave school because their needs are not being met in the educational setting (Birdsall & Correa, 2007; Hansen & Toso, 2007; Matthews, 2006). It is important to recognize that when gifted students drop out, the system has failed them and their potential is wasted. Matthews recommended teachers watch for students who show attendance, discipline, or academic problems as these are sometimes early markers for students at risk for dropping out. Ford and colleagues (2007) extended the data to provide suggestions for teachers and parents to use as positive pressure for low income gifted students to remain in school and achieve.

Girls.

Noble (1987), Reis (2003), Roeper (2003), and Subotnik and Arnold (1993) are names synonymous with the study of gifted girls. Through their research and publications, they have helped form a solid foundation for others to build upon in the field of gifted girls. Kerr (1997) worked to “discover the barriers to gifted females’ achievement and investigative means of overcoming them” (p. 483). Noble and Subotnik (1999) and Subotnik and Arnold (1995) reported on the identification of gifted girls. Kerr noted that “an early start may be critical to gifted girls’ receiving an appropriate education,” (p. 489). Additionally, Kerr suggested selecting tests without bias and that the “multidimensional approach is most useful in selecting gifted girls for broad-based programs of enrichment or other increased academic challenges” (p. 490).

Rebhorn and Miles (1999) also expressed concerns about the disadvantages gifted girls encounter in high-stakes testing like the SAT. However, by 2002 Dai had refuted those studies indicating that current research shows the gender gap closing. Furthermore, Schober and colleagues (2004) and Lubinski, Benbow, Shea, Efekhari-Sanjani, and Halvorson (2001) found the topic of the gender gap to be obsolete as well.

Nevertheless, Nelson and Smith (2001) and Roeper (2003) found that although gifted girls may have overcome those societal notions and beaten the gender gap, they are still underrepresented in high-status, typically male-dominated fields. Both studies indicated that environment (home, peers, school, etc.) was a primary contributing factor to those results (Nelson & Smith; Roeper). Both Arnold (1993) and Noble (1987) pointed out some discrimination and struggles remain for gifted young women in college and as they enter the workforce.

Ironically, after decades of published work about the disadvantages girls overcome in our current school system, recently there has been a surge of publicity regarding the disadvantages boys face in the public school arena. Sadker (2002) suggested that perhaps the most meaningful debate regarding gender differences in schools is to acknowledge that school has never been a utopia and that some population will inevitably struggle.

The focus on sub-populations must be clarified, as it is not an effort to discriminate or single out, but to identify unmet areas. While great strides have been made to identify gifted students from minority populations, work remains to be done. There is still a marginalized segment of students who are rarely identified for gifted programs – those who are already identified as a special education student. Bianco (2005) found that neither general education nor

special education teachers were able to accurately identify students for gifted programs when a special education label was present. These are hurdles that must be constantly challenged. Additionally, almost all high stakes testing includes a component of investigating the sub-populations in given districts and schools to ensure that students are achieving beyond the labels.

Four Tests Used in Local School District

This study evaluated the efficacy of four specific aptitude and achievement tests used to identify gifted learners for a middle school English course. The participating Texas school district offers a pull-out enrichment program for elementary students. In middle and high school, gifted students are grouped homogenously in English and math courses. There are specific entrance criteria for each type of program: different instrument combinations are used for the elementary, English, and math programs. Once admitted to the middle/high school program, students must maintain a nine-week grade average of 80% or above. If a student's average is below 80% for two consecutive nine weeks, he or she can be dismissed from the program.

This study focused exclusively on the criteria for admittance to the middle school gifted English program and the efficacy of those assessments to select successful students for the program. The four tests used were the Cognitive Abilities Test, the Iowa Test of Basic Skills, the STEP Writing test, and the Naglieri Nonverbal Ability Test.

Cognitive Abilities Test (CogAT).

The Cognitive Abilities Test (CogAT) is a group-administered abilities test intended to measure mental abilities developed by Lohman and Hagen (2001). The most recent norming data

were collected in 2005. This up-to-date comparison of students' abilities in reasoning and problem solving is commonly used as a measure for determining aptitude.

The CogAT is comprised of three separate sections and resulting scores: verbal, non-verbal, and quantitative. The CogAT provides an intelligence quotient standardized with other aptitude tests; a mean score of 100 and standard deviations of 15 points each. As is customary in most states, "gifted" is defined as students who score two standard deviations above the norm – 130 or above. It was included in this study as one of the four test scores used to determine eligibility for the gifted English middle school program. Having been in existence and used widely for many years, it is considered valid and reliable (Lohman & Hagen).

Iowa Test of Basic Skills (ITBS).

The Iowa Test of Basic Skills (ITBS) (Hoover et al., 2001) is a group-administered achievement test intended to provide a comprehensive assessment of student progress in major content areas. The ITBS includes a series of subtests that measure achievement in vocabulary, word analysis, reading comprehension, listening, language, mathematics, social studies, science, and in the upper levels only, sources of information. The norms were most recently conducted in 2005, with the copyright dates ranging from 2001 to 2007, depending on the form used. Test–retest stability coefficients over a one year time interval were mostly in the .70 to .90 range and internal consistency and alternate forms reliability coefficients were in the .80s and .90s (Gregory, 1996). Additionally, reviews have been very positive, noting sound measurement practices, high technical standards (i.e., internal consistency and alternate forms reliability), and good content validity (Canivez, 2000).

Widely used in school districts across the nation, the ITBS is a commonly used assessment to mark student achievement nationally – individually and collectively (Canivez, 2000). Students earning scores at the 95th percentile or above are regularly considered advanced. This assessment instrument has been part of numerous studies because of its sound reliability and validity (Bracey, 2009; Jenner, & Jenner, 2007; Kim & White, 2008; Schilling, Carlisle, Scott, & Ji Zeng, 2007). The ITBS Reading test is also one of the four instruments used to assess students' eligibility for the gifted English middle school program.

STEP Writing.

Evaluating writing abilities is a time-consuming and subjective task; therefore, the STEP Writing Test, as a multiple-choice instrument used to objectively evaluate editing and revising abilities of students, is a powerful tool (Madaus & Rippey, 1966). Anderson wrote an article in 1960 introducing the STEP Writing Test. The only other research using this instrument was published in 1966, when Madaus and Rippey studied the 1957 version to determine its validity and address concerns regarding a multiple-choice measurement to be used for writing evaluation.

For the current study, extensive research was done across available databases (such as the Academic Search Complete, Dissertation and Theses A&I, Professional Development Collection, ERIC, JSTOR, and Google Scholar), but the out-of-print test has not appeared in the research since. Though the Madaus and Rippey study (1966) was older, it was one of the most extensive examinations of the comparison of writing skills and content with an objective-style writing test. The two-year study involved high school English teachers, who taught structured writing assignments, and required four writing samples (Madaus & Rippey). “The total writing score of the STEP does not seem to be related strongly to any of the individually important

writing behaviors our English teachers felt minimally necessary for good composition” (Madaus & Rippey, p. 22); however, this study “did not attempt to validate this global aspect of STEP but only attempted to see how STEP compared with seven limited aspects of actual writing behavior” (p. 22). Since STEP provides the teacher with information of how a student stands in relation to a norm group on some kind of global writing score” (p. 22), it functions well as part of a larger assessment protocol. There is value in understanding the nature of writing and how the STEP can provide discrete, accurate data that can be collected and demonstrated efficiently to establish minimal foundational writing skills.

The STEP Writing test was used in this study as part of the combined instruments to evaluate the students’ understanding of writing norms. These results are used to ascertain how qualified a student might be for this advanced English course. Because the test was no longer in print and its norming data were out of date, the district program director recognized the need to update the norms information when she joined the district in 1996. Therefore, according to the program director, the school district has renormed the STEP test more than once in the past 14 years to ensure that it is valid and reliable for the local population and purpose.

In order to conduct the norming process in a fair and appropriate way, a committee was formed of highly experienced and well-trained teachers in the gifted program to conduct the norming. The program director served as statistician and directed all norming procedures. That process included gathering STEP Writing test results from all fifth-grade students, plus data from students who completed the test in the prior three years. The committee reconvened to evaluate the score distribution and adjust the cut-off score accordingly. The program director reported that a similar committee will be formed for the purpose of renorming this assessment again in the fall

of 2012. Even though it is out of print, the evaluation committee finds the results from the STEP Writing test to be valuable, reliable, valid, and an excellent component in the collection of instruments used to evaluate prospective students for the English gifted middle school program according to the program director.

Naglieri Nonverbal Ability Test (NNAT).

Within the field of gifted education, there has been a push to find ways to eliminate test bias and constructs that might impede identification of students who might not traditionally qualify because of race, gender, experiences or socio-economic status (Lohman, 2006). To that end, researchers have explored the use of non-verbal assessments (Lohman, 2005b). The Naglieri Nonverbal Ability Test (NNAT) (Naglieri, 2003) is one such instrument. Nasca (1988) investigated “the effect of using a non-verbal measure of intellectually functioning to supplement more traditional measures of I.Q.” (p. 1). Traditionally, the Wechsler Intelligence Scale for Children-Revised (Wechsler, 2003), the Slosson, and the Otis-Lennon (Otis & Lennon, 1997) had been used in the school system of interest (Nasca).

Nasca (1988) studied the validity of incorporating two nonverbal measures; the Test of Non-Verbal Intelligence (Brown, Sherbenou, & Johnsen, 2010) and Raven’s Progressive Matrices (Raven, Court, & Raven, 1983) in conjunction with traditional measures. Both nonverbal tests “require inductive reasoning grounded in geometric and figural relationships” (Nasca, p. 1). In an effort to include students that might otherwise not qualify with more traditional measures of intelligence, the use of non-verbal tests potentially affected 6.9%-28.6% of the studied population (Nasca). Correlations were lower than expected, perhaps “due to the unique nature of populations” (p. 8), but the study found the data could be used “to determine the

impact of various decision making strategies, e.g., minimum cut-off scores, % of the population based on weighted scores, minimum score on any one measure, etc.” (Nasca, p. 9). In the conclusions, Nasca warned that “specific recommendations of the use of non-verbal measures of intellectual functioning will become available only after the predictive validity of identification criteria has been established” and that until then “only qualified generalizations are available” (p. 9). Others have followed up Nasca’s research to validate instruments and study effectiveness with minority and English language learner populations (Frasier, 1997; Hunsaker et al., 1991; Lohman; Lohman et al., 2008).

The test has seven levels, each consisting of 38 items. Adequate reliability is suggested because the internal consistency estimates reliability range from .83 to .93 across grades. However, the corresponding standard errors of measurement (SEM) of the Nonverbal Ability Index (NAI) score ($M = 100$, $SD = 15$) range from 5.6 to 5.8 (Naglieri, 1997). The 95% confidence interval for NAI scores of 2 SEM is 22.8 or about 1.5 *SD*.

According to the multilevel norms booklet (1997), “the Naglieri Nonverbal Ability Test is a brief, culture-fair, nonverbal measure of school ability” (p. 5). Like other nonverbal tests, the NNAT requires students to use problem-solving and reasoning skills to answer questions in an effort to determine aptitude. “The NNAT is an extension and revision of the Matrix Analogies Test ... used extensively in educational and research settings” (p. 5). The use of this instrument as part of the entrance criteria helps to provide additional information, particularly for students who might benefit from a wordless test to assess their aptitude. Following a recommendation from the state education agency, the school district in this study adopted the NNAT as part of its entrance assessments for students at all levels.

Summary

This literature review provided a foundation of relevant, timely, appropriate, and essential research for this dissertation. Since there is no single definition of intelligence or gifted, each school district is allowed to interpret the federal definition and determine programming. While intelligence and achievement have both been measured and debated for about a century, gifted education has really only been in the forefront for a few decades. Despite the efforts across many years and many theorists, there is a lack of agreement on the definition and measurement of intelligence, as well as the definition and measurement of giftedness. Even within the field of gifted education, there has been specific curiosity about the quality of the instruments we choose to select students for gifted programs (Benbow, 1992; Gustin & Corazza, 1994; Kerr, 1997; Lohman, 2005a; Lohman et al., 2008; Niederer et al., 2003). With these discrepancies of philosophies, individual state and district program directors are called upon to not only decide an appropriate definition and assessment for giftedness, but also determine a matching program (Lohman; Louis et al., 2000; Nettelbeck & Wilson, 2005; Samuels, 2007; Sternberg & Clinkenbeard, 1995). This dissertation study sought to determine the efficacy of the combined identification assessments and the relationship with student earned grades for a specific school in central Texas.

Chapter Three

Methodology and Procedures

This chapter explains the research methodology employed in this study. The first section describes the purpose of the study and delineates the research design. The second section includes the research questions and null hypothesis. The third section defines the participant selection process and procedures. This chapter concludes with the data collection, explanation of the instruments used, the data analysis procedures, and the limitations of the study.

Purpose of the Study

This study investigated the relationship between the combination of four entrance assessments and student performance in a Texas middle school academically gifted English language arts program. The intent was to evaluate the efficacy of the combined instruments used to identify students who were successful in the program. Individual success was determined by the grade each student earned in the gifted middle school English class, as that is the criterion the school district uses to dismiss unsuccessful students. Students' with grades below 80% for a nine-week grading period were placed on probation; two consecutive nine-week periods with below an 80% average in the academically gifted English middle school class could result in removal from the program.

Research Design

This was a quantitative research study that used a bivariate Pearson correlation study to investigate the relationship between entrance assessment scores and student grades for a middle school gifted English course. Several regression models, including an analysis of repeated measures, was used because students were measured on four different entrance assessments (not always given in the same order to all students) and on their nine-week averages over the course of more than one year. This also required the use of a Multivariate Analysis of Variance (MANOVA) to evaluate whether the scores varied across multiple tests and multiple nine-week periods.

The data of incoming sixth graders were collected longitudinally for the years they were in the program. The study was designed to include the data for these students while taught by a single teacher. However, due to financial constraints within the district, programming changes occurred at the middle school mid-way through the study, beginning in the fall of 2010. The gifted courses were moved from a single teacher for all three years of the program to different teachers in each grade level. So, students who entered the gifted English middle school program in the fall of 2007 as sixth graders have scores for all three years they were in the program. Students who entered in the fall of 2008 as sixth graders have scores for their sixth and seventh grade years. Finally, because the selected teacher remained assigned as the seventh-grade instructor in the program, the students who entered the program in the fall of 2009 as sixth graders have scores for their sixth and seventh grade years as well.

When conducting correlation studies, it is important to be aware of factors affecting the interpretation of results. In this specific case, examples might be subjective grading, motivation

of student/underachievement, and life events like divorce, moves, and deaths. Those limitations were recognized and taken into account when data were analyzed by examining the data for outliers and other anomalies. Outliers might occur if a student neglected school when his or her family was in turmoil; other students might have sought attention by purposefully failing a class. These would be considered outliers pragmatically because their performance does not reflect whether or not they were capable of doing the work in the gifted program (Hinkle, Wiersma, & Jurs, 1998). Additional outliers may have occurred due to some unusual admission practices the director of gifted programming may have employed from time to time. For example, when some students scored above the 95th percentile on only two instruments, the director often made a judgment call about whether the student would likely find success in the program. Also, there are instances when the director of gifted programming may have admitted students to the program because of parent protest, teacher recommendation, or other circumstances. These instances were also taken into account when outliers were found in the data analysis. It is also important to recognize that when conducting repeated measures studies some of the factors affecting the interpretation of repeated measures (e.g., time of day, room temperature, behavior of test administrator, order of the tests given).

Research Questions

The following research questions were designed to structure the study:

1. Is there a correlation between the percentile scores earned on the battery of tests given and the middle school gifted English student grades?
2. To what extent did the battery of tests used as entrance criteria for the middle school gifted English program positively correlate with student grades? Did the battery of tests

and decisions based on those scores serve as statistically significant predictors of student success?

3. Which entrance criteria most strongly predict academic success in the middle school gifted English program?

Null Hypotheses

1. There is no correlation between the evaluation measures used to qualify students and student grades for a middle school gifted English program.
2. The scores on the combined entrance tests do not relate to student grades.
3. Higher scores on the entrance measures do not relate to higher student grades.
4. No single entrance measure predicts academic success in the middle school English program better than the combination of the four instruments.

Participant Selection

Archival data were used from the spring semester testing of 2007 through the fall of 2009. These longitudinal data included the nine-week averages of students in sixth grade for the 2007-2008 school year, nine-week averages for students in sixth and seventh grades for the 2008-2009 school year, nine-week averages for students in all three grade levels for the 2009-2010 school year, and nine-week averages for students in seventh grade for the 2010-2011 school year. Each student was enrolled in the program no less than one complete nine-week grading period.

The total population of students in the district who qualify for academically gifted English middle school program averages 1,500 students for most years. Participants for this

study were selected from a single middle school with approximately 150 students in the academically gifted middle school English program spanning four consecutive years. This was a convenience sample (Hinkle, Wiersma, & Jurs, 1998). Specifically, using only one middle school does not allow one to generalize the results to the entire population of middle school students in this district, but it did provide sufficient confidence to conduct a study to determine if further study was warranted district-wide. The population of the entire middle school was approximately 1,250 students. The ethnic composition of the student population was African American (2.9 %), Asian (2.6 %), Caucasian (48.2%), Hispanic (42.6 %), and other (3.7%). The ethnic composition of the study group was African American (2%), Caucasian (70%), Hispanic (26%), and other (2%). The participants' school opened in 1982 in a suburban area.

This study investigated the efficacy of the entrance measures for the academically gifted middle school English program in a large suburban Texas district. The district has 13 middle schools, each providing its own gifted English program. However, due to variances in the size of the program across the district and the principals' discretion, the program varies widely. For example, at some schools where the gifted population is small, several grade levels are combined or gifted students are combined with advanced students. On other campuses, there are enough gifted students to provide at least one section of gifted English in addition to the regular and advanced English courses offered to the other students on each grade level. In most of these schools, different teachers are responsible for each grade level. In only a few schools is the entire gifted population taught by a single teacher.

The 150 participants for this study were students in a single middle school gifted English program in the district. These students were taught all years of middle school gifted English (6th,

7th and/or 8th grades) by a single teacher, thereby minimizing the variance of experiences, course material, teacher-constructed assessments, instructional setting, and opportunities for earning grades. While it might imply that student performance was confounded with the teacher, maintaining a single teacher for this study was selected in an effort to minimize the teacher variances that might contribute to student success. The students were given the entrance assessment between March 2006 and August 2009 by either the district program director or another qualified individual. All participating students were enrolled in the gifted middle school English program at least one semester during the 2007-2008, 2008-2009, 2009-2010, or 2010-2011 school years.

The sample was selected not only due to convenience but also because it helped control many of the variables that might determine variations in student grades. For instance, student grades may be highly subjective and may vary from teacher to teacher (Robins & Jolly, 2011; Stanley & Baines, 2004). By constricting the sample to a single teacher, the designed experiment minimized that subjectivity. Additionally, by limiting participants to those from a single campus, other factors that might affect the outcomes were minimized (e.g., school climate regarding intelligence and achievement, size of gifted program and its impact on school climate, hardships faced by some campuses due to student deaths or violence on campus, teacher rapport with students).

The factors of a common instructor and environment led to a more homogeneous group than one might obtain by looking at the entire population of gifted middle school students in gifted English programs across multiple campuses. However, these same constraints also limit

the predictive nature of the expected results; one cannot extend or generalize the results obtained from this study to other schools or other districts.

At this campus, classes are approximately 45 minutes long. This teacher combined a wide variety of means to convey information and conducted her class during these 45 minutes (e.g., lecture, partner/group work, individual work, discussions, presentations, student-led lessons). Each student attended eight assigned classes each day, including gifted English. Grades were earned by completing daily assignments (e.g., homework, class work, quizzes) and assessments (e.g., vocabulary tests, unit tests, major papers, projects, presentations). Student grades were on a modified 10-point scale, with the range of 100-90 an A, 89-80 a B, 79-75 a C, 74-70 a D, and anything below 69 considered an F. Students earned a 0 if no work was completed on an assignment, and points were deducted from work that was submitted late.

The data used in the study are considered archival, and none of the students in the study is taught by this teacher any longer. The middle school grades and entrance scores are no longer relevant or have any impact on their participation in the gifted program. The University of New Orleans Committee for the Protection of Human Subjects in Research determined that the research and procedures were exempt from federal regulations under 45 CFR 46.101 Category 1 (Appendix C).

Procedures

In compliance with the Human Subjects Review Committee guidelines, this study was conducted under the supervision and approval of the University of New Orleans Graduate School. At the local level, permission to conduct research with student data was secured from the

superintendent's office (Appendix D), the director of the gifted program (Appendix E), and the middle school principal (Appendix F). As per the Human Subjects requirements, all records, materials, and data collected were hand-carried from the data-collection location and maintained by the researcher, and all forms and codes identifying a particular student were destroyed immediately following data analysis.

Data Collection

The data were collected by the researcher and gifted program director for the district, and then coded to ensure anonymity for all students. The program director provided some or all of the following data for each student: student-coded number, gender, percentile score for the Cognitive Abilities Test (CoGAT) (Lohman & Hagen, 2001); percentile score for the Iowa Test of Basic Skills: Reading (ITBS) (Hoover et al., 2001); percentile score for the STEP Writing (Educational Testing Services, 1957); and percentile score for the Naglieri Nonverbal Abilities Test (NNAT) (Naglieri, 2003). The campus data processor and teacher provided nine-week grades for each eligible year of admission and the date students were removed from the program, if applicable.

The data were analyzed using Statistical Package for Social Sciences (SPSS) software. A bivariate correlation study was used to evaluate the strength of the relationship between each of the four entrance tests and student grades earned in the middle school gifted English course. Both linear and logistic regression models were run, as well as a repeated-measures regression study, to determine which instrument was most effective in predicting success in the gifted English middle school program.

Instruments

The four entrance assessments used to determine eligibility for the gifted middle school English program were the CogAT, ITBS Reading, STEP Writing, and NNAT. Each was given by the district gifted education program coordinator or a trained individual. The testing rooms were typically at the students' school – though some were tested during the summer in the central office building – but all sites were determined to be test-friendly and comfortable by the district gifted educator program coordinator. The testing was given over a two-day time period in most cases. Exceptions typically occurred when a student enrolled in the district from out of the area, and they already had some of the test scores and did not need to take all tests for admittance to the program.

Cognitive Abilities Test (CogAT).

The Cognitive Abilities Test (CogAT) is a group administered abilities test intended to measure mental abilities. Developed by Lohman and Hagen (2001), it is considered valid and reliable (Lohman, 2003). The most recent norming data were collected in 2005. The internal consistency is between .85-.98 (Robins & Jolly, 2011). The CogAT compares students' reasoning and problem-solving abilities and is comprised of three separate sections and resulting scores: verbal, nonverbal, and quantitative. The CogAT provides an intelligence quotient standardized with other aptitude tests, a mean score of 100 and standard deviations of 15 points. As is customary in most states (NAGC, 2011), "academically gifted" is defined as students who have a superior intellect or advanced learning ability, often recognized as an aptitude score two standard deviations above the norm – 130 or above.

Iowa Test of Basic Skills (ITBS).

The Iowa Test of Basic Skills (ITBS) (Hoover et al., 2001) is a group-administered achievement test intended to provide a comprehensive assessment of student progress in major content areas. The ITBS includes a series of subtests that measure achievement in vocabulary, word analysis, reading comprehension, listening, language, mathematics, social studies, science and, in the upper levels only, sources of information. The norms were most recently collected in 2005, with the copyright dates ranging from 2001 to 2007, depending on the form used. Test-retest stability coefficients over a one-year time interval were mostly in the .70 to .90 range, and internal consistency and alternate forms reliability coefficients were in the .80s and .90s (Gregory, 1996). Additionally, reviews have been very positive, noting sound measurement practices, high technical standards (i.e., internal consistency and alternate forms reliability), and good content validity (Canivez, 2000).

Widely used in school districts across the nation, the ITBS is a commonly used norm-referenced assessment to mark student achievement nationally – individually and collectively (Canivez, 2000). Students earning scores at the 95th percentile or above are regularly considered advanced. This assessment has been part of numerous studies because of its demonstrated reliability and validity (Bracey, 2009; Jenner, & Jenner, 2007; Kim & White, 2008; Schilling et al., 2007). Only the results of the Reading section of the test were included in this study's assessments given to students seeking admission to the academically gifted English middle school program.

STEP Writing.

The STEP Writing test (Educational Testing Service, 1957) was introduced when Anderson (1960) wrote an article about it. After an exhaustive search in multiple databases, it was found that the only other research using this instrument was published in 1966, when Madaus and Rippey studied the STEP Writing test to determine its validity and addressed concerns regarding a multiple-choice measurement to be used for writing evaluation. Since the “STEP provides the teacher with information of how a student stands in relation to a norm group on some kind of global writing score” (p. 22), it functions well as part of a larger assessment protocol. There is value in understanding the nature of writing and how the STEP can provide discreet, accurate data that can be collected and demonstrated efficiently to establish minimal foundational writing skills (Madaus & Rippey). Due to its age, information regarding its initial reliability and validity is scarce; this is a noted limitation of this study.

The STEP Writing test was used in this dissertation study as part of the combined instruments to evaluate the students’ understanding of writing norms. These results are used to ascertain how qualified a student might be for the targeted gifted English course. Because the test is no longer in print and its norming data that are out of date, the district program director for gifted education recognized the need to update the norms information when she joined the district in 1996. Therefore, this local Texas school district has renormed the STEP test more than once in the past 16 years to ensure that it is valid and reliable for the specific population and purpose locally. It was not renormed during the study period (2007-2011).

In order to conduct the norming process in a fair and appropriate way, a committee was formed of highly experienced and well-trained teachers in the gifted program to conduct the

norming. The program director served as statistician and directed all norming procedures. That process included gathering STEP Writing test results from all fifth-grade students, plus data from students who had completed the test in the prior three years. The committee reconvened to evaluate the score distribution and adjust the cut-off score accordingly. The program director reported that a similar committee will be formed for the purpose of renorming this assessment again in the fall of 2012. Even though it is out of print, the district evaluation committee found the results from the STEP Writing test to be valuable, reliable, valid, and a component in the collection of instruments used to evaluate prospective students for the English gifted middle school program. The researcher did not have access to this evaluation and recognizes the associated limitations with this instrument.

Naglieri Nonverbal Ability Test (NNAT).

According to the multilevel norms booklet (Naglieri, 1997), “the Naglieri Nonverbal Ability Test is a brief, culture-fair, nonverbal measure of school ability” (p. 5). Like other nonverbal tests, the NNAT (Naglieri, 2003) requires students to use problem-solving and reasoning skills to answer questions in an effort to determine aptitude. “The NNAT is an extension and revision of the Matrix Analogies Test ... used extensively in educational and research settings” (p. 5). The use of this instrument as part of the entrance criteria helps to provide additional information, particularly for students who might benefit from a wordless test to assess their aptitude. Following a recommendation from the state education agency, the Texas school district in this study adopted the NNAT as part of its entrance assessments for students at all levels.

The test has seven levels, each consisting of 38 items. Adequate reliability is suggested because the internal consistency estimates reliability range from .83 to .93 across grades.

However, the corresponding SEM of the Nonverbal Ability Index NAI score ($M = 100$, $SD = 15$) range from 5.6 to 5.8 (Naglieri, 1997). The 95% confidence interval for NAI scores of 2 SEM is 22.8 or about 1.5 SD.

The scores were used as individual predictors in the study. Each student's four entrance scores were compared with their nine week grades. The individual assessment scores were kept to determine if cases of outliers had any implications for students whose grades indicated any academic struggle.

Data Analysis

The data analysis included correlation analyses of the variables using Pearson product-moment correlation coefficients and multiple-regression analyses. The data were analyzed to determine r . Once the correlation data were analyzed, a repeated-measures regression study was designed to establish which instrument(s) predicted course averages most strongly. Eleven tables are used to graphically display the data and results.

Limitations

The most notable limitation in this study is the use of percentile scores for data analysis. Though widely used, percentile scores are inherently non uniform; equal differences in percentiles do not represent equal differences in raw scores (Hinkle et al., 1998). Such scores are not useful when comparing different assessments. Furthermore, the use of percentile scores tend to exaggerate small, nearly nonexistent differences at the center of a normal distribution, whereas percentile scores tend to underestimate actual differences in the tails of the distribution (Hinkle et al.). However, the LEA gifted program director did not record raw scores that could be converted

to standard scores for analysis; instead, the only data provided to the researcher were the percentile scores. This is a significant limitation of the study and will affect the strength of the results. However, there is still potentially valuable information that can be ascertained for this specific situation.

Another limitation of the study was the expected range restriction of the data (Goodwin & Leech, 2006; Stringer, 2008). Because students already scored well enough on the entrance assessments (averaging 95% and above) to qualify for the gifted program, the range or variability of scores for comparison was relatively narrow. And, because the grades examined only included those of students who had scored well on the entrance assessments and rarely fall below the 80% cut-off, the narrow range produced a relatively low correlation. This is a by-product of studying a very homogeneous group in a very homogeneous setting. This range restriction tends to narrow the level of statistically significant difference in the results.

A further limitation was the fact that the process and matrix used by the LEA to evaluate student scores on entrance assessments were statistically unsound. The lack of consideration for standard error of measurement was another concern. Additionally, the divisions of scores on the matrix appeared to be arbitrary and not research based.

Another statistical limitation stemmed from the convenience sample, as the sample selection bias may have undermined the external validity of the study. Additionally, this convenience sample did not represent the experience of students across the district; however, it was effective to determine how district-wide further investigation could be conducted.

An additional limitation centered on the classroom teacher. Since all of the students were taught by a single instructor, student performance could be confounded with instruction. Further concerns associated with this limitation include potential instructor bias, grading subjectivity, and teacher expectations for specific products or projects. Selecting a single instructor produced consistency and limited some of these biases, but there is no way to control for all of them. Consequently, this study potentially measured the teaching ability of the instructor instead of the efficacy of the entrance instruments.

The STEP Writing test was also addressed as a serious limitation in this study. The renorming of this instrument may not have been conducted with appropriate statistical protocols. Initially, this instrument was not intended for a gifted audience. Additionally, the publication date and lack of research regarding the test brought to light some concerns; however, it was concerns such as these, in part, that prompted the study. The STEP Writing test had been used for many years in this school district and one outcome of this study was to determine its effectiveness as a tool for the middle school English gifted program.

Finally, that the ability to generalize these results was restricted due to some of these limitations. This means that further forms of analysis or a follow-up study may be warranted.

Summary

The purpose of this study was to evaluate the efficacy of the combined instruments used to identify students who were successful in the program by determining the strength of the relationship between the combined entrance criteria for a middle school gifted English program and the grades students earned in that course. The four instruments, CogAT, ITBS, STEP

Writing, and NNAT, were compared to the nine-week averages of each enrolled student. This quantitative study used a bivariate Pearson correlation study and MANOVA to evaluate the relationships.

Archived data (entrance scores and nine-week averages) were collected for all the students enrolled in the middle school English gifted program at a single school, taught by the same teacher for all years in the study. There were positive and negative aspects of this convenience sample, but it ultimately served to determine that the entire district middle school English gifted program entrance criteria needs to be evaluated.

Participants' identities were protected and all requirements of the University Committee for the Protection of Human Subjects in Research were met, including obtaining an exemption from federal regulations under 45 CFR 46.101 Category 1.

There were numerous statistical limitations including issues with percentile scores, range restriction, the convenience sample, and generalizability. Additional limitations include the concerns regarding the STEP Writing test and the limitation of using a single classroom teacher. However, the study is a warranted first step to evaluating the efficacy of the entrance criteria for this Texas middle school academically gifted English program.

Chapter Four

Results

The purpose of this study was to determine if the entrance assessments given to students entering a middle school academically gifted English course were accurate predictors of students who would be successful in the program (as measured by their grades each nine-week grading period). The four entrance assessments evaluated were the Cognitive Abilities Test (CogAT), Naglieri Nonverbal Ability Test (NNAT), STEP Writing, and the Iowa Test of Basic Skills: Reading (ITBS). The data were analyzed using a variety of statistical analysis (e.g., correlation, linear regression, logistic regression) in order to answer the research questions guiding this study:

1. Is there a correlation between the percentile scores earned on the battery of tests given and the middle school gifted English student grades?
2. To what extent did the battery of tests used as entrance criteria for the middle school gifted English program positively correlate with student grades? Did the battery of tests and decisions based on those scores serve as statistically significant predictors of student success?
3. Which entrance criteria most strongly predict academic success in the middle school gifted English program?

Four null hypotheses were tested.

Null Hypothesis #1: There is no correlation between the evaluation measures used to qualify students and student grades for a middle school gifted English program.

Null Hypothesis #2: The scores on the combined entrance tests do not relate to student grades.

Null Hypothesis #3: Higher scores on the entrance measures do not relate to higher student grades.

Null Hypothesis #4: No single entrance measure predicts academic success in the middle school English program better than the combination of the four instruments.

The results of the data analysis rejected all the null hypotheses. Following a demographic description of the participants and a section describing the methods and procedures used, the findings will be presented in terms of the research questions.

Demographic Description

Participants were selected from a large suburban Texas school district (65,000 students). The total population of students in the district who qualify for gifted English middle school program averages 1,500 students in most years. From the 13 middle schools, a convenience sample (Hinkle et al., 1998) of approximately 150 students was selected from a single middle school. Additionally, a single teacher taught sixth -through eighth-grade gifted English for several years, and the data collected were only from the years they were taught by this teacher.

The overall student population of the participating middle school was approximately 1,250 students. The ethnic composition of the population was African American (2.9 %), Asian

(2.6 %), Caucasian (48.2%), Hispanic (42.6 %), and other (3.7%). The ethnic composition of the study group was African American (2%), Caucasian (70%), Hispanic (26%), and other (2%).

The mean age for most sixth graders is 12 in the study. The school opened in 1982 in a suburban area.

The group of students that entered the program for the 2007-2008 school year as sixth graders remained in the study for all three years of middle school (6th, 7th, and 8th grades). There were 60 students in this group, 28 females (47%) and 32 males (53%). The second group of students entered the program for the 2008-2009 school year as sixth graders and remained in the study for two consecutive years (6th and 7th grades). There were 53 students in this group, 26 females (49%) and 27 males (51%). The third group of students that entered the program as sixth graders for the 2009-2010 school year also remained in the study for two consecutive years (6th and 7th grades). There were 38 students in this group, 16 females (42%) and 22 males (58%). Data for this information may be found in Table 1.

Table 1

Demographic Information and Participation Years Display

	Gender		2007-2008	2008-2009	2009-2010	2010-2011
	M	F				
Cohort 1	32	28	6 th grade	7 th grade	8 th grade	
<i>N = 60</i>	53%	47%				
Cohort 2	27	26		6 th grade	7 th grade	
<i>N = 53</i>	51%	49%				
Cohort 3	22	16			6 th grade	7 th grade
<i>N = 38</i>	58%	42%				

Methods and Procedures

The number and nature of the research questions demanded a variety of statistical analysis. Correlation analysis, logistic regression and a generalized linear model were used. Normality is not a consideration when logistic regression models are used because there is only a yes/no type of determination.

Prior to analysis, the data were reorganized so that each student's nine week grade, with their associated entrance assessment percentile scores, became a single item for analysis. Most of the first group of students were in the program for all three years, and therefore, had 12 records (one for each nine-week grading period). Data were collected for students in the second and third groups for only two years, yielding eight records for most of them. This enabled each student to have equal weight in the analysis as it was organized by nine-week grades rather than completed semesters or years. Data for these results may be found in Table 2. These reorganized data generated 1,200 discrete records for analysis. However, since some students did not have all four entrance assessment scores, not all were included in analysis. There were 28 missing cases (mostly ITBS Reading achievement scores), so the scores of 1,172 (97.7%) of the participants' were analyzed.

Table 2

Total Sample Summary

Unweighted Cases ^a		<i>N</i>	Percent
Selected Cases	Included in Analysis	1172	97.7
	Missing Cases	28	2.3
	Total	1200	100.0
Unselected Cases		0	.0
Total		1200	100.0

As a first step, a binary logistic regression model was run to determine if students who were dismissed from the program were potentially capable of success based on their entrance assessment scores. Data from these results are presented in Table 3. Students who were successful were coded with a binary code of 1 ($n = 1,111$); unsuccessful students were coded as 0 ($n = 61$). Of the 1172 students who could be included in the analysis, the logistic regression would classify all students as successful based on the logistic regression model. The model correctly classified all successful students as successful but likewise classified all unsuccessful students as successful as well. The data revealed that students' entrance scores did not indicate that any student might not be successful. In other words, the current entrance assessments accurately find students capable of success in the middle school gifted English program in the participating district. Since some students are not successful, however, other investigations were warranted. The data presented in Table 4 contain the coefficients and associated p -values of this logistic regression model. In this regression, only the percentiles from the battery of tests were used.

Table 3

Prediction of Success

Observed			Predicted		
			Success		Percentage Correct
			0	1	
Step 0	Success	0	0	61	.0
		1	0	1,111	100.0
		Overall Percentage			94.8

Note. Constant is included in the model. The cut value is .500

Table 4

Prediction of Success Coefficients

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a Gender	.515	.328	2.458	1	.117	1.673
Grade	.273	.205	1.774	1	.183	1.313
Quarter	.167	.124	1.816	1	.178	1.181
CogAT	-.134	.076	3.075	1	.079	.875
NNAT	-.127	.036	12.731	1	.000	.881
ITBS Reading	.031	.015	4.298	1	.038	1.031
STEP	.077	.021	14.187	1	.000	1.080
Constant	15.445	8.092	3.643	1	.056	5099617.820

^a. Variable(s) entered on Step 1: Gender, Grade, Quarter, Aptitude, NNAT, Reading, STEP.

Research Question #1: Is there a correlation between the percentile scores earned on the battery of tests given and the middle school gifted English student grades?

To answer this question, a Pearson correlation r was chosen to examine correlations among the variables – the four entrance assessments (CogAT, ITBS, Step Writing, and NNAT). These results are presented in Table 5. The N varies because some students were not given all four assessments. This occurred if a student had similar test scores from another district or if a student was also testing for the academic gifted math program the same day since the district gifted coordinator is aware that students taking too many tests in a short period of time may not perform well on all measures.

The low correlation numbers demonstrated that the four assessments may sample different content or ability areas, with negative correlations between CogAT Aptitude with both the ITBS Reading (-.110) and STEP Writing tests (-.124). The NNAT might be positively correlated with the CogAT (.159). This is possible since both measure aptitude. However, it is negatively correlated with the ITBS Reading (-.080); these are similar to the results of the CogAT Aptitude test. The ITBS Reading and STEP Writing tests were positively correlated with one another (.081), as both measure achievement in the language arts area. However, the STEP Writing seemed to be positively correlated with the NNAT (.023) and the ITBS Reading (-.080) was not.

The correlation analysis may have demonstrated that each of the four entrance tests assess different content or ability. It may also have demonstrated that the aptitude tests (CogAT and NNAT) are not confounded with the ITBS Reading achievement or STEP Writing scores for this study, as they are negatively correlated. They may all measure cognitive abilities or subject-

matter knowledge, but they are not sampling the same behaviors. This finding supports the researcher's premise that these assessments may measure different aspects of cognitive abilities and experiences. While significant in some cases, all of the relationships were small.

Table 5

Four Entrance Assessments Correlation

		CogAT	NNAT	ITBS Reading	STEP
CogAT	Pearson Correlation	1	.159**	-.110**	-.124**
	Sig. (2-tailed)		.000	.000	.000
	<i>N</i>	1188	1188	1172	1188
NNAT	Pearson Correlation	.159**	1	-.080**	.023
	Sig. (2-tailed)	.000		.006	.419
	<i>N</i>	1188	1200	1184	1200
ITBS Reading	Pearson Correlation	-.110**	-.080**	1	.081**
	Sig. (2-tailed)	.000	.006		.006
	<i>N</i>	1172	1184	1184	1184
STEP	Pearson Correlation	-.124**	.023	.081**	1
	Sig. (2-tailed)	.000	.419	.006	
	<i>N</i>	1188	1200	1184	1200

**Significant at the 0.01 level (2-tailed).

An additional regression analysis of the scores (grades) compared with the scores from the battery of four tests was used to answer this first research question. These results are presented in Table 6. The R^2 was just over 8% with the adjusted R^2 just below that, so it was not

a statistically significant result, in general. However, the ITBS Reading achievement score ($p = .001$) and the STEP Writing test score ($p = .000$) appear to be the best predictors of students' grades. These findings indicated that of the four entrance assessments given, the ones best able to predict student grades were the ITBS Reading Achievement score and the STEP Writing test. Neither the CogAT Aptitude test ($p = .301$) nor the NNAT ($p = 1.07$) seems to be a significant predictor of student success in the program.

Table 6

Regression Analysis of Scores Against Tests

Model	Unstandardized Coefficients		Standardized Coefficients	<i>t</i>	Sig.
	B	Std. Error	Beta		
1 (Constant)	65.387	6.172		10.595	.000
CogAT	-.049	.047	-.030	-1.034	.301
NNAT	-.028	.017	-.046	-1.612	.107
ITBS Reading	.079	.023	.098	3.439	.001
STEP	.272	.031	.248	8.721	.000

Note. Dependent variable: Score

Research Question #2: To what extent did the battery of tests used as entrance criteria for the middle school gifted English program positively correlate with student grades? Did the battery of tests and decisions based on those scores serve as statistically significant predictors of student success?

A second logistic regression model was run similar to the one used to answer the first research question, but in this case the predictors were augmented with cofactors of gender, grade level, quarter (nine-week grade), in addition to the scores obtained on CogAT, NNAT, ITBS Reading, and STEP. These results are presented in Table 7.

As illustrated, the ITBS Reading and STEP Writing test predicted grades with p -values of 0.038 and 0.000, respectively. This multiplicative model also revealed that the NNAT was inversely related to grades ($p < 0.000$) whereas CogAT was not a significant predictor of grades ($p < 0.079$). Moreover, the cofactors of gender, grade, and quarter were not significant. These findings were surprising, as the researcher expected the STEP Writing test to be ineffective and outdated.

Table 7

Logistic Regression: Reading and STEP Significance

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1						
Gender	.515	.328	2.458	1	.117	1.673
Grade	.273	.205	1.774	1	.183	1.313
Quarter	.167	.124	1.816	1	.178	1.181
CogAT	-.134	.076	3.075	1	.079	.875
NNAT	-.127	.036	12.731	1	.000	.881
ITBS Reading	.031	.015	4.298	1	.038	1.031
STEP	.077	.021	14.187	1	.000	1.080
Constant	15.445	8.092	3.643	1	.056	5099617.820

Research Question #3: Which entrance criteria most strongly predict academic success in the middle school gifted English program as measured by nine-week grades?

A generalized linear model was used to estimate students' score (not simply student success as previously discussed) by using the fixed factors of gender, grade level, and grades as covariates of the four entrance assessments. These results are presented in Table 8.

This repeated-measure ANOVA analyzed within-groups comparisons such as entrance scores and grades earned in the course. The effect of the grades with the STEP Writing test had an F value of 66.154 with one degree of freedom. The effect of the grades with the ITBS Reading achievement score had an F value of 9.461 with one degree of freedom. It is significant at the .002 level. The scores on the CogAT aptitude test ($p < .923$) and the NNAT ($p < .530$) were statistically insignificant. Finally, the predictive power of the model was poor, as both the raw R^2 and the adjusted R^2 were just above the 12% level. Even when insignificant terms were discarded from the model (i.e., NNAT and Aptitude), the R^2 from this analysis is about 11%. There was not enough evidence to support a linear relationship between the predictors and student grades and no test predicts student grade with this analysis.

Table 8

Repeated Measure ANOVA Without ID: Reading and STEP Significance

Source	Type III Sum of Squares	df	Mean Square	<i>F</i>	Sig.
Corrected Model	5652.777 ^a	14	403.770	12.523	.000
Intercept	2838.242	1	2838.242	88.029	.000
Year	480.778	3	160.259	4.971	.002
Gender	800.877	1	800.877	24.840	.000
Grade	351.097	3	117.032	3.630	.013
Quarter	383.477	3	127.826	3.965	.008
Aptitude	.305	1	.305	.009	.923
NNAT	12.708	1	12.708	.394	.530
Reading	305.038	1	305.038	9.461	.002
STEP	2132.949	1	2132.949	66.154	.000
Error	37304.031	1157	32.242		
Total	9745559.000	1172			
Corrected Total	42956.808	1171			

^a. *R* squared = .132 (adjusted *R* squared = .121).

Further results indicated that as the students remained in the program, their grades increased. These results are presented in Table 9. This might indicate that unsuccessful students were removed earlier, that students learned to achieve for that specific teacher over time, or that the curriculum became easier. An additional logistic regression was run including ID of the

individual student. Data for these results are presented in Table 10. The most significant predictor was the individual student. The R^2 was elevated to 72.9% and the adjusted R^2 was 68.9%. Only grade level was not significant at the 5% level. A second analysis was made using the generalized linear model, using year, grade, quarter, and individual student. The individual student was the most significant predictor. The R^2 was elevated to 72.9% and the adjusted R^2 was 68.9%. Grade level was not significant at the 5% level, indicating that students did not uniformly perform better with a specific grade-level curriculum than another.

Table 9

Nine-week Grades and Averages

	6 th Grade				7 th Grade				8 th Grade			
	1 st	2 nd	3 rd	4 th	1 st	2 nd	3 rd	4 th	1 st	2 nd	3 rd	4 th
Cohort 1: Enter 2007	93	91	91	91	88	92	89	93	93	90	93	92
Average	92				91				92			
Cohort 2: Enter 2008	89	87	88	91	91	89	90	92				
Average	89				91							
Cohort 3: Enter 2009	91	92	89	92	92	92	91	91				
Average	91				92							

Table 10

Repeated Measure ANOVA With ID: Individual Student Significance

Source	Type III Sum of Squares	df	Mean Square	<i>F</i>	Sig.
Corrected Model	31308.779 ^a	150	208.725	18.296	.000
Intercept	1165.079	1	1165.079	102.124	.000
Year	121.974	3	40.658	3.564	.014
Gender	.000	0	.	.	.
Grade	73.041	3	24.347	2.134	.094
Quarter	329.775	3	109.925	9.635	.000
CogAT	.000	0	.	.	.
NNAT	.000	0	.	.	.
ITBS Reading	.000	0	.	.	.
STEP	.000	0	.	.	.
ID	25656.001	136	188.647	16.536	.000
Error	11648.029	1021	11.408		
Total	9745559.000	1172			
Corrected Total	42956.808	1171			

^a. *R* squared = .729 (adjusted *R* squared = .689).

An additional logistic regression analysis was run to determine if gender was a factor in student success. Data for these results are presented in Table 11. When all other factors are included in the model, the coefficient is -0.03, indicating that the gender factor for males is only

97% of that for females. This means that females earned slightly higher grades than males in the academically gifted English middle school program during this time frame.

A final analysis used regression linear estimation. These results are presented in Table 12. It was run with score (grade) as the dependent variable against the predictors of student ID and the four entrance assessments. The R^2 was 86% and the adjusted R^2 was 82%. This analysis revealed that there was a significant difference in individual students, as revealed by a significance of .020. Neither aptitude (CogAT) (.407) nor NNAT (.216) was significant, but the reading (ITBS) (.000) and STEP (.000) were significant. The two aptitude tests (CogAT and NNAT) were not good predictors of student success in the program, but the two achievement instruments (ITBS Reading and STEP Writing test) seemed to better predict which students will be successful.

Table 11

Logistic Regression: Gender

	Score	df	Sig.
Step 0 Variables Gender(1)	13.074	1	.000
Grade	4.723	2	.094
Grade(1)	4.723	1	.030
Grade(2)	2.314	1	.128
Aptitude	14.097	1	.000
NNAT	18.367	1	.000
Reading	11.794	1	.001
STEP	22.726	1	.000
Overall Statistics	59.784	7	.000

Table 12

Linear Estimation: Prediction of Success

Model	Unstandardized Coefficients		Standardized Coefficients	<i>t</i>	Sig.
	B	Std. Error	Beta		
1 (Constant)	63.753	6.200		10.283	.000
ID	-.009	.004	-.067	-2.333	.020
CogAT	-.039	.048	-.024	-.830	.407
NNAT	-.022	.018	-.036	-1.237	.216
ITBS Reading	.080	.023	.100	3.522	.000
STEP	.278	.031	.253	8.904	.000

^a. Dependent variable: Score

Chapter Five

Conclusions

Summary of the Findings

The purpose of this study was to investigate the relationship between the entrance assessments for a middle school gifted English program and the grades students earned once placed in that program. The school district policy included dismissing students if they did not maintain the required 80% course average. Because some students did not find success by maintaining this minimum average after being admitted to the academic gifted program, the study investigated the entrance assessments to determine if they were appropriately identifying students for the program. The data included entrance scores and student-earned grades for two or three years for each of the 150 participants. Each student's entrance assessment scores were paired with a nine-week grade for each nine-week period that a grade was recorded - this yielded 1,200 discrete records for data analysis.

The quantitative methodology used in this study included correlation and regression models. Using logistic regression and generalized linear regression models, the repeated measures were analyzed for their predictive values. SPSS software was used for all data analyses.

While the exact data findings are detailed in Chapter Four, the following is a rich discussion of the data analysis. The discussion includes a separate section detailing the

implications for gifted programs, particularly in the participating LEA. The final sections include recommendations for further study and a summary.

Findings

The following is a summary of the findings revealed through data analysis:

1. The results indicated that of the four entrance assessments given, the ones most able to predict student grades were the ITBS Reading Achievement score and the STEP Writing test.
2. Neither the CogAT Aptitude test nor the NNAT (Naglieri Nonverbal Ability Test) was a significant predictor of student success in the program.
3. As students remained in the program, their nine-week grades increased.
4. When all other factors held constant, females' grades were predicted to be about two points higher than those of males.
5. The most significant predictor was the individual student.

The results rejected all the null hypotheses. Additionally, the results indicated that students selected for participation in the program were all capable of success, whether they remained in the program or not. Further findings revealed that students did not uniformly perform better with a specific grade level curriculum than another, meaning that the curriculum appears to be equally rigorous for each year of the program.

A caution must be reiterated regarding the use of the percentile scores for data analysis. The missing raw score data created an imperfect study and further evaluation was necessary, it

was deemed an acceptable flaw. Readers must be aware that the results could be skewed due to this type of data usage and may vary when this study is repeated with standard scores.

Discussion

This discussion section covers five topics: achievement and aptitude, grades, teacher effectiveness, gender, and individual students. The first section, covering achievement and aptitude, addresses the first two findings. The third finding is addressed in the sections on grades and teacher effectiveness. The fourth and fifth findings are addressed in the sections covering gender and individual student, respectively.

Aptitude and achievement.

The results showed the two aptitude tests (CogAT and NNAT) were not good predictors of student success in the program, but the two achievement instruments (ITBS Reading and STEP Writing test) seemed to better predict which students were successful. Two aspects need clarification here. First and foremost, in this study the statistical results indicated that these four specific tests do not measure the same abilities; the aptitude tests and achievement tests were not positively correlated with one another. In fact, both aptitude measures (i.e., CogAT and the NNAT) had an inverse relationship with ITBS Reading; the CogAT with the STEP Writing test. This was an issue of concern at the outset of this study, and it is a highly debated issue in the field. Thus, some theorists argue this potential interrelatedness can cause concern when both aptitude and achievement measures are used to make judgments about students, while others like Assouline (1997) assert that there are differences between what aptitude and achievement tests measure.

Second, a distinction must be made between identification of gifted students and prediction of student success, as measured by grades. There are dual purposes at work here, so clarification is necessary. Identifying students who are gifted generally requires an assessment of their aptitude. This is a foundational concept for many gifted programs (“exhibits high performance capability in an intellectual...area” as defined by Texas Education Code, 2009). Aptitude tests like the ones used in this study propose to measure aptitude; indeed, a century of research has revealed that the concept of *g* is widely useful in school settings (Gottfredson, 2004). Quality gifted programs typically include aptitude tests in the student evaluation component to determine if the student is gifted (Robinson et al., 2000). However, this study was not just investigating whether students were accurately identified as gifted but also whether those entrance assessments were accurately predicting student success. Once a student has established superior mental capacity through aptitude assessment, his or her success in a rigorous academic program can be assessed. Data analysis illustrates that formally measured student achievement predicted student achievement in this academic middle school gifted English program.

Grades.

In the same way, a previous grade earned in an English language arts class might be useful for identifying successful students for this academically gifted middle school English program. Caution is warranted, however, as grades are not comparable from teacher to teacher (Greville, 2009; Robins & Jolly, 2011; Stanley & Baines, 2004) and may be the result of teacher expectations for high achievers to continue to achieve (Siegle & Reis, 1998). Nevertheless, the director of the gifted education program revealed that previous English language arts grades are sometimes collected. Such grades have frequently (and anecdotally) been found to be an

important piece of evidence to predict future success and should be used in future student qualifying evaluations (Pestello, 1987; Willingham, Pollack, & Lewis, 2002).

A related finding was that students' grades increased over time. There are several possible reasons for these results. One explanation is that student grades increased from year to year because unsuccessful students were removed from the program in previous years; however, the percentage of students removed from the program was relatively low (ranging from 2% - 8%). Another possible explanation is that the curriculum became easier. Other possible factors explaining the increase in student grades include teacher expectations, increasing student self-efficacy, building core knowledge and skills needed in subsequent years, that these achieving students learned to achieve better for this particular teacher, or some combination of these possibilities.

Teacher effectiveness.

A final comment about student grades must be made because the study was designed to include only a single teacher so as to minimize the variance of experiences, course material, teacher-constructed assessments, instructional setting and opportunities for earning grades. A question that emerged during data analysis was whether the study, in part, was measuring the effectiveness of the curriculum, instruction, or this particular teacher instead of the entrance criteria. This teacher, like others in the district assigned to teach classes for gifted students met the state minimum requirement of 30 hours of teacher inservice training. Additionally, this teacher routinely attended a six-hour training meeting each fall and sometimes an additional six-hour training meeting in the spring. Common topics addressed at those training meetings include differentiation within the gifted class, depth and breadth of the gifted curriculum, common

struggles of gifted students, characteristics of a gifted classroom, and characteristics of an effective gifted teacher, among others. It is through these meetings that the district director for gifted education strives to ensure quality instruction in every gifted class in the district. However, participation in those trainings does not ensure an effective gifted teacher. Furthermore, it does not discriminate between student performance, motivation, and quality instruction. This factor may need to be considered in follow-up studies since there is no way to know the depth and breadth of the curriculum, or the level of critical thinking or problem solving this particular teacher required of her gifted students.

Gender.

Another interesting result from the data analysis was that when all other factors were held constant, females' class grades were predicted to be about two points higher than those of their male counterparts. The fact that females potentially would score better than males may be an effect of puberty and differing rates of maturation, as females tend to mature at earlier ages than males (Spencer, Dupree, Swanson, & Cunningham, 1998). Thus, Spenser and colleagues noted positive and negative social and psychological effects, particularly during middle school years, between early and late-maturing males and females. Freeman (2004), as well as Tang and Neber (2008), noted cultural factors that might influence gender and achievement. Further influential factors may be student attitudes (Swiatek & Lupkowski-Shoplik, 2000) or teacher perceptions (Siegle & Reis, 1998), as the teacher is a female.

Programming match.

Another concern raised earlier in this study was the match between the entrance assessments and the program for which they are serving as gatekeeper. Louis and colleagues

(2000) emphasize that assessments ought to identify students who will profit from placement in the program. This concern was allayed because the results indicated that there was a match between some of the assessments and the program. Given that the ITBS: Reading and STEP Writing achievement tests were both found to have a positive relationship with student grades, these assessments are accurately being used to identify students who will find success in the academically gifted middle school English program.

However, another aspect of program match should be mentioned. While these assessments match the program for which they are selecting students, there are still students for whom this academically gifted middle school English program is not a suitable match. In as much as this study attempted to determine if those assessments were successful in selecting students for the program, it is important to note that there was not a match for some students. Each student must be considered individually, especially when success is not found.

Individual students.

Another significant predictor of student success was the individual student, as measured by the combination of their entrance scores and grades. It was established in the first analysis that all students were correctly identified as gifted and as being capable of success in the program. Any number of triggers can impact student achievement (e.g. life events, motivation, rapport with the teacher).

Since students were unsuccessful, it is important to explore possible causes. Some of those causes are related to the student's home life, peer pressures, societal expectations, and motivation. Because underachievement among gifted students has plagued educators for many years, it was considered as a potentially significant contribution to this study. Numerous

scientific writings, books, articles, and programs have been designed to identify and address the specific needs of this subpopulation of hidden and often unidentified gifted students (Delisle, 2006; Delisle & Galbraith, 2002; Neihart, 2006; Peterson, 2002; Rimm, 1997b).

Implications for Gifted Programs

The first recommendation for the LEA upon which this study was based is to collect raw scores on all future student entrance evaluations to expedite future program evaluation. The missing data impaired this study but can easily be rectified by recording the raw score with the percentile score earned by each potential student.

The second recommendation for the LEA is to increase the qualitative assessments in the initial screening and identification of potentially gifted students, as both quantitative and qualitative data should be considered (NAGC, 2010b; Ryser, 2011). The NAGC Pre-K-Grade 12 Gifted Programming Standards provide a foundation of appropriate guidelines to use for identification, placement, and programs for gifted students. NAGC and Ryser both asserted that further use of teacher and parent input might prove helpful as well, as anecdotal and other qualitative information. Since teacher and parent input is not a required element in the assessment of students for this middle school gifted English program, the district might be missing students who do not perform well on tests or are atypical in their presentation of giftedness. (e.g., minority students, students with disabilities). While the district has a plan in place for identifying underrepresented gifted students, Hispanic students have been underrepresented in the program at this particular campus during these years.

Further investigation may be warranted to determine if that is an on-going issue at this campus and/or a district wide underrepresentation. Countless sources of evidence demonstrate

that having additional qualitative information would amplify the probability of locating and servicing underrepresented students (Borland & Wright, 1994; NAGC, 2010b; Ryser, 2011). This may expand the nomination and evaluation net used to identify potentially gifted students, thereby selecting more students for the program that might not otherwise be chosen.

A related recommendation involves exploring and finding more authentic ways to assess the success of students in the academically gifted English middle school program. Grades are highly subjective and, therefore, not considered authentic assessment for all students. Numerous alternatives could replace or augment the current evaluation program for students in the academically gifted middle school English program. “Alternative assessment tools like performance-based assessment and portfolios are more useful in gauging the authentic level of performance” in gifted students (Van Tassel-Baska, 2001, p. 67).

Additionally, anecdotal evidence suggests that previous achievement in a nongifted English language arts program might be a good indicator for future success. Collecting the previous class grade may help indicate students who will not find success without intervention for possible underachievement. In fact, Bowers (2009) concurred that previously earned grades can be beneficial tools for assessing student performance. Certainly caution should be exercised as varying curriculum, instructors, and level of rigor may play heavily into the lack of usefulness of those scores.

Furthermore, a team of gifted educators in the district needs to investigate the basis for the dismissal and probation protocols in the district, evaluating whether they are research-based, effective, and appropriate.

Since the results might suggest that motivation and student underachievement are possible causes of unsuccessful students in the academic gifted program, the LEA needs to devote some resources to prepare the teachers in the gifted program to recognize students at-risk for underachievement. The NAGC (2010b) Pre-K-Grade 12 Gifted Programming Standards may be used to evaluate related topics such as classroom environment, curriculum planning and instruction, and professional development. Underachievement, as measured in this study by students unable to maintain the required minimum average, is considered a typical struggle among gifted students (Delisle & Galbraith, 2002; Morisano & Shore, 2010; Rimm, 2012; Siegle, 2012), and indeed all students. Delisle (2006) preferred the term “selective consumers,” inferring that students perform when it is of interest or relevance to them (p. 97). Current research points to evidence that numerous students are victims of underachievement. For example, Strip and Hirsch (2000) recommended an initial evaluation to rule out vision, hearing or other physical or emotional health issues, and then beginning to investigate the issues behind underachievement. It is a complex problem with many contributing factors (Rimm, 1996). Consequently, there are numerous specialists and published works (Delisle; Peterson, 2002; Rimm) on the topic. Several in the field suggest interventions to help these students (Emerick, 1992; Frey, 2002; Morisano & Shore; Rimm, 1997a; Siegle), including goal-setting, organizational assistance, and forming an alliance between the student, parents, and teacher.

Just as this study brought to light the possibility that the data may measure teacher effectiveness as much as the efficacy and predictive value of the entrance assessments, it also questions the validity of the curriculum used in the district. For example, one might question if the individual teacher included sufficient higher level questions, critical thinking, or appropriately structured activities to reach exclusively gifted students, instead of merely

advanced learners who were also high achievers. One might also question if the curriculum aligned more with the STEP Writing test and ITBS Reading test because it is not a rigorous gifted program expecting students to perform at levels more commensurate with the CogAT and NNAT.

A recommendation to the director for gifted education is to form a team to evaluate the current curriculum in the gifted program against the national standards as defined in the NAGC (2010b) Pre-K-Grade 12 Gifted Programming Standards. This ensures that student performance is a measurement of their abilities and achievement instead of the suitability of the curriculum.

Recommendations for Further Study

The first recommendation must be to repeat the present study using standard scores derived from the raw scores of each entrance assessment since those scores were unavailable for this study. Evaluating the district wide program in the near future is a distinct possibility as long as the raw scores are available. This would confirm or refute findings in the current study and, more importantly, contribute to the research field of gifted education.

A related recommendation is to collect demographic information for each campus gifted program. During the years of this study, the middle school's demographics changed due to a sharp increase in the number of students qualifying for free and reduced-price lunch, a common indicator of socio-economic cultural influences. This may have affected the number of students in the gifted middle school program or the school climate towards giftedness.

Even though the STEP Writing test's ability to predict success in the academic gifted English classroom was found to be significant, its lack of contemporary national norming data

and studies related to its reliability and validity are still issues of concern. Experts caution the use of instruments that have not been nationally normed more recently than 14 years (Robins & Jolly, 2011). Further study could include evaluating the STEP Writing test to find an appropriate replacement or another suitable assessment.

Additionally, further study could explore optional instruments and ways to evaluate student writing abilities or look to find a more recently developed measure that would provide similar information, perhaps the Test of Written Language-2 (TOWL-2) (Hammill & Larsen, 1988). The district could also consider creating a professionally suitable instrument for local usage or renorming the STEP Writing test nationally.

Evidence suggests that there is also a need to investigate the effectiveness of different gifted teachers, and a study to that effect should be conducted in this district. Since the program varies from campus to campus within the district, both campus-based and district wide evaluation might prove beneficial. One consideration is the ever-changing fluctuation of middle school gifted teachers, as principals often rotate that responsibility. It would be unfortunate to study the effectiveness of a particular teacher on a specific campus, only for that teacher to be reassigned elsewhere. Hence, an overall assessment of the entire district might provide the most useful results. If student motivation, as measured by students who do not maintain the required minimum grade average, appears as a dominant factor of student success, then training for personnel to recognize and intervene for students at risk for underachievement would be essential.

Finally, a future study could implement specific interventions with students at risk of underachievement to see if it affects their performance. Such interventions could include some or

all of the six steps in Rimm's Trifocal Model (2012), modeling curiosity (Siegle, 2012), or goal-setting (Morisano & Shore, 2010), to name a few. Bembenutty (2009) argued that expectations of grades should not be the sole focus of students' motivation. These results would be a serious contribution to research in the field of gifted education (Neihart, 2006; Reis & McCoach, 2000).

Summary

The purpose of this chapter was to share and discuss the results from the data analysis. The findings were presented as five statements that encapsulated the outcomes and were organized into five divisions: achievement and aptitude, grades, teacher effectiveness, gender, and individual student. These discussions amplified the data results and explained the repercussions of the findings.

The next section was a discussion regarding the implications of gifted programming, particularly for this local district. Specifically, this included recommendations for the director of gifted education to apply within the school district, including teacher training, assessment selection, and curriculum evaluation. These interventions cover a wide range of topics to provide the director of gifted education with a variety of avenues to explore evaluation for the program. This academically gifted program was determined to be of high quality, but could still improve in several areas.

The final sections detailed four possible recommendations for further study. While other studies may be generated from this (e.g., district wide evaluation of similar data, effect of puberty on academic achievement for gifted adolescents), these three were the most widely applicable and would serve to fill gaps in the research literature. These included repeating the study using standard scores instead of percentile scores, determining if the STEP Writing test

could be replaced or renormed nationally, evaluating teacher effectiveness, and incorporating interventions for underachievement.

At the outset of this research study, there were some expected results. Being able to predict student success is a long-studied and highly practical skill for educators. Researchers and educators alike are interested in finding which standardized scores can provide this data. This study found that two of the four instruments used to evaluate students for the academically gifted middle school English program may do just this. However, the significance of the individual student and their related motivation was an interesting insight. For this school district, the individual teachers need to be able to differentiate and properly motivate the atypical gifted learner. The results suggest that the most influential skill-set an educator of gifted adolescents may need are those related to identifying and servicing students who are at risk of underachievement.

References

- Adelson, J. L. (2007). A “perfect” case study: Perfectionism in academically talented fourth graders. *Gifted Child Today*, 30(4), 14-20.
- Amend, E. R., Schuler, P., Beaver-Gavin, K., & Beights, R. (2009). A unique challenge: Sorting out the differences between giftedness and Asperger’s Disorder. *Gifted Child Today*, 32(4), 57-63.
- Anderson, C. C. (1960). The new step essay test as a measure of composition ability. *Educational and Psychological Measurement*, 20(1), 95-102. doi: 10.1177/001316446002000110
- Arnold, K. D. (1993). Undergraduate aspirations and career outcomes of academically talented women: A discriminant analysis. *Roeper Review*, 15(3), 169-175.
- Assouline, S. G. (1997). Assessment of gifted children. In N. Colangelo & G.A. Davis (Eds.), *Handbook of gifted education* (2nd ed., pp. 89-108). Boston, MA: Allyn and Bacon.
- Assouline, S. G., & Lupkowski-Shoplik, A. (1997). Talent searches: A model for the discovery and development of academic talent. In N. Colangelo & G.A. Davis (Eds.). *Handbook of gifted education* (2nd ed., pp. 89-108). Boston, MA: Allyn and Bacon.
- Baker, B. D. (2001). Gifted children in the current policy and fiscal context of public education:

- A national snapshot and state-level equity analysis of Texas. *Educational Evaluation and Policy Analysis*, 23(3), 229-250.
- Bembenuddy, H. (2009). Teaching effectiveness, course evaluation, and academic performance: The role of academic delay of gratification. *Journal of Advanced Academics*, 20(2), 326-355.
- Benbow, C. P. (1992). Academic achievement in mathematics and science of students between ages 13 and 23: Are there differences among students in the top one percent of mathematical ability? *Journal of Exceptional Psychology*, 84, 51-61.
- Berlin, J. E. (2009). It's all a matter of perspective: Student perceptions on the impact of being labeled gifted and talented. *Roeper Review*, 31, 217-223. doi: 10.1080/12783190903177580
- Betts, G. T. (1985, November). *Autonomous learner model for the gifted and talented*. Paper presented at the 32nd Annual Convention of the National Association for Gifted Children. Denver, CO.
- Bianco, M. (2005). The effects of disability labels on special education and general education teachers' referrals for programs. *Learning Disability Quarterly*, 28(4), 285-293.
- Birdsall, P., & Correa, L. (2007). Gifted underachievers. *Leadership*, 36(4), 21-23.
- Bollag, B. (2005). English teachers blast SAT writing test. *Chronicles of Higher Education*,

51(36).

Bonner II, F. A. (2000). African American giftedness: Our nation's deferred dream. *Journal of Black Studies, 30*(5), 643-663.

Borland, J. H., & Wright, L. (1994). Identifying young potentially gifted, economically disadvantaged students. *Gifted Child Quarterly, 38*, 164-171.

Bowers, A. J. (2009). Reconsidering grades as data for decision making: More than just academic knowledge. *Journal of Educational Administration, 47*(5), 609-629.

Bracey, G. W. (2009). A miracle on 125th street? *Phi Delta Kappan, 91*(1), 91-92.

Bradley, G. (2001). Is the SAT falling out of favor? *Academe, 87*(4), 14-16.

Bridgeman, B., Burton, N., & Pollack, J. (2008). Predicting grades in college courses: A comparison of multiple regression and percent succeeding approaches. *Journal of College Admission, 199*, 19-25.

Brody, L. E. (2005). The study of exceptional talent. *High Ability Studies, 16*(1).

Bronson, P. (2009, September 18). In defense of the SAT [Web log post]. Retrieved from <http://blog.newsweek.com/blogs/nurtureshock/archive/2009/09/18/in-defense-of-the-sat.aspx>

Brown, L., Sherbenou, R. J., & Johnsen, S. (2010). *Test of Nonverbal Intelligence: A Language*

Free Measure of Cognitive Ability. Texas: Pro-Ed.

Burris, C. C., & Welner, K. G. (2005). Closing the achievement gap by detracking. *The Phi Delta Kappan*, 86(8), 594-598.

Büttner, G., & Hasselhorn, M. (2011). Learning disabilities: Debates on definitions, causes, subtypes, and responses. *International Journal of Disability, Development & Education*, 58(1), 75-87. doi: 10.1080/1034912X.2011.548476

California Achievement Test, Fifth Edition (CAT/5). (2009). Monterey, CA: CTB Mc Graw-Hill.

Camara, W., & Echternacht, G. (2000). *The SAT I and high school grades: Utility in predicting success in college* (College Board Report No. RN-10). New York, NY: College Entrance Examination Board.

Canivez, G. L. (2000). Predictive and construct validity of the developing cognitive abilities test: Relations with the Iowa test of basic skills. *Psychology in the Schools*, 37(2), 107-113.

Carr, D. (2009). Revisiting the liberal and vocational dimensions of university education. *British Journal of Educational Studies*, 57(1), 1-17. doi: 10.1111/j.1467-8527.2009.00425.x

Christodoulou, J. A. (2009). Applying multiple intelligences. *School Administrator*, 66(2), 22-26.

Clark, B. (1997). *Growing up gifted: Developing the potential of children at home and at school* (5th ed.) Upper Saddle River, NJ: Merrill.

Colangelo, N. (1997). Counseling gifted students: Issues and practices. In N. Colangelo & G.A.

Davis (Eds.). *Handbook of gifted education* (2nd ed., pp 353-365). Boston, MA: Allyn and Bacon.

Colangelo, N., & Davis, G. A., (Ed.). (1997). *Handbook of gifted education*. (2nd ed.).

Boston, MA: Allyn and Bacon.

Colangelo, N., Assouline, S. G., Marron, M. A., Castellano, J. A., Clinkenbeard, P. R., Rogers,

K., Calvert, E., Malek, R., & Smith, D. (2010). Guidelines for developing an academic acceleration policy. *Journal of Advanced Academics*, 21(2), 180-203.

Coleman, M. R., & Galagher, J. J. (1995). State identification policies: Gifted students from

special populations. *Roeper Review*, 17(4), 268-275

Coleman, M. R., Harradine, C., & King, E. (2005). Meeting the needs of students who are twice

exceptional. *Teaching Exceptional Children*, 38(1), 5-6.

Coughlin, D. (1997). The person with a learning disability as a minority group member. *Journal*

of Learning Disabilities, 30(5), 572-575.

Cross, T. L. (2009a). Disrupting social contracts that affect gifted students: An homage to

Harrison Bergeron. *Gifted Child Today*, 32(4), 14-15.

Cross, T. L. (2009b). Social and emotional development of gifted children: Straight talk. *Gifted*

Child Today, 32(2), 40-41, 65.

Cross, T. L., & Coleman, L. (1993). The social cognition of gifted adolescents: An exploration of the Stigma of Giftedness Paradigm. *Roeper Review*, 16(1), 37-40.

Dai, D. Y. (2002). Are gifted girls motivationally disadvantaged? Review, reflection, and redirection. *Journal for the Education of the Gifted*, 25(4), 315-358.

Dai, D. Y., Swanson, J. A., & Cheng, H. (2011). State of research on giftedness and gifted education: A survey of empirical studies published during 1998-2010 (April). *Gifted Child Quarterly*, 55(2), 126-138. doi: 10.1177/0016986210397831

Delisle, J. (1986). Death with honors: Suicide among gifted adolescents. *Journal of Counseling & Development*, 64(9), 558-561.

Delisle, J. (1999). A millennial hourglass. *Gifted Child Today*, 22(6), 26-32.

Delisle, J. (2003). To be or to do: Is a gifted child born or developed? *Roeper Review*, 26(1), 12-13.

Delisle, J., & Galbraith, J. (2002). *When Gifted Kids Don't Have All the Answers*. Minneapolis, MN: Free Spirit.

Delisle, J. R. (2006). *Parenting gifted kids: Tips for raising happy and successful children*. Waco, TX: Prufrock Press.

Doina, Jr., R. D. (1997). Evaluation of gifted programs. *Gifted Child Today Magazine*, 20(5), 38-40.

Dweck, C. S. (2010). Even geniuses. *Educational Leadership*, 68(1), 16-20.

Educational Testing Service. (1957). *Sequential tests of educational progress manual for interpreting scores: Writing*. Princeton, NJ: Author.

Emerick, L. J. (1992). Academic underachievement among the gifted: Students' perceptions of factors that reverse the pattern. *Gifted Child Quarterly*, 36(3), 140-146.

Erford, B. T., Vitali, G. J., & Slosson, S. W. (1999). *Manual for the Slosson Intelligence Test-Primary (SIT-P)*. East Aurora, NY: Slosson Educational.

Feldhusen, J. F. (1997). Educating teachers for work with talented youth. In N. Colangelo & G. A. Davis (Eds.). *Handbook of gifted education* (2nd ed., pp. 547-552). Boston, MA: Allyn and Bacon.

Feldman, D. H. (1993). Child prodigies: A distinctive form of giftedness. *Gifted Child Quarterly*, 37(4), 188-194.

Fletcher, J. M., Denton, C., & Francis, D. J. (2005). Validity of alternative approaches for the identification of learning disabilities: Operationalizing unexpected underachievement. *Journal of Learning Disabilities*, 38(6), 545-552.

Ford, D. Y. (1995). Desegregating gifted education: A need unmet. *The Journal of Negro Education, 64*(1), 52-62.

Ford, D. Y. (2004). A challenge for culturally diverse families of gifted children: Forced choices between achievement and affiliation. *Gifted Child Today, 27*(3), 26-27.

Ford, D. Y., & Grantham, T. C. (2003). Providing access for culturally diverse gifted students: From deficit to dynamic thinking. *Theory into Practice, 42*(3), 217-225.

Ford, D. Y., Grantham, T. C., & Frazier-Trotman, M. (2007). Diamonds in the rough: Identifying and serving low income students. *Understanding Our Gifted, 19*(3), 12-14.

Ford, D. Y., Grantham, T. C., & Whiting, G. W. (2008). Another look at the achievement gap: Learning from the experiences of gifted black students. *Urban Education, 43*(2), 216-239.

Fraas, J. W., & Frick, C. (1992). *Can the Kindergarten Diagnostic Instrument be used to identify children who will qualify for a reading intervention program?* (Research Report).
Lexington, OH: Lexington Local School District. (ED349549)

Francis, D. J., Fletcher, J. M., Stuebing, K. K., & Lyon, G. R. (2005). Psychometric approaches to the identification of LD: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities, 38*(2), 98-108.

- Frasier, M. M. (1997). Gifted minority students: Reframing approaches to their identification and education. In N. Colangelo & G. A. Davis (Eds.). *Handbook of gifted education* (2nd ed., pp. 498-515). Boston, MA: Allyn and Bacon.
- Freeman, J. (2004). Cultural influences of gifted gender achievement. *High Ability Studies*, 15(1), 7-24. doi: 10.1080/1359813042000225311
- Frey, C. (2002). *Dealing with the needs of underachieving gifted students in a suburban school district: What works!* Retrieved from <http://www.gifted.uconn.edu/nrcgt/newsletter/spring02/sprng023.html>
- Gagne, F. (2004). Transforming gifts into talents: The DMGT as a developmental theory. *High Ability Studies*, 15(2), 119-147. doi: 10.1080/1359813042000314682
- Galton, F. (1869). *Hereditary genius*. New York, NY: Macmillan.
- Gardner, H. (1983). *Frames of mind*. New York, NY: Basic Books.
- Gehring, J. (2002). Amid criticism, College Board considers revamping SAT. *Education Week*, 21(29), 16-17.
- Geiser, S., & Studley, R. (2002). UC and the SAT: Predictive validity and differential impact of the SAT I and SAT II at the University of California. *Educational Assessment*, 8(1), 1-26.
- Goodwin, L. D., & Leech, N. L. (2006). Understanding correlation: Factors that affect the size of

r. *The Journal of Experimental Education*, 74(3), 251-266.

Gordon, S., Duff, S., Davidson, T., & Whitaker, S. (2010). Comparison of the WAIS-III and WISC-IV in 16-year-old special education students. *Journal of Applied Research in Intellectual Disabilities*, 23(2), 197-200. doi: 10.1111/j.1468-3148.2009.00538.x

Gottfredson, L. S. (2004). Schools and the g factor. *Wilson Quarterly*, 28(3), 35-45.

Gregory, R. J. (1996). *Psychological testing: History, principles, and applications* (2nd ed.). Boston, MA: Allyn & Bacon.

Greville, E. C. (2009). A rose by any other name: Grading and assessment. *Assessment Update*, 21(5), 1-2.

Grobman, J. (2006). Underachievement in exceptionally gifted adolescents and young adults: A psychiatrist's view. *The Journal of Secondary Gifted Education*, 17(4), 199-210.

Grossman, F. M., & Johnson, K. M. (1983). Validity of the Slosson and Otis-Lennon in predicting achievement of gifted students. *Educational and Psychological Measurement*, 43, 617-622.

Guilford, J. P. (1982). Cognitive psychology's ambiguities: Some suggested remedies. *Psychological Review*, 89, 48-95.

Gustin, W., & Corazza, L. (1994). Mathematical and verbal reasoning as predictors

of science achievement. *Roepers Review*, 16(3), 160.

Hammill, D. D., & Larsen, S. C. (1988). *Test of Written Language-2*. Austin, TX: Pro-Ed.

Hansen, J., & Linden, K. (1990). Selecting instruments for identifying gifted and talented students. *Roepers Review*, 13(1), 10.

Hansen, J., & Toso, S. J. (2007). Gifted dropouts: Personality, family, social, and school factors. *Gifted Child Today*, 30(4), 30-41.

Hargrove, K. (2008). Meeting social and emotional needs in the days of high-stakes testing. *Gifted Child Today*, 31(3), 45-46.

Herrnstein, R. J., & Murray, C. A. (1994). *The bell curve: Intelligence and the class structure in American life*. New York, NY: Free Press.

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1998). *Applied statistics for the behavioral sciences*. Boston, MA: Houghton Mifflin.

Hollingworth, L. S. (1938). Bright students take care of themselves. *The Modern Language Journal*, 22(5), 385.

Hoover, H. D., Dunbar, S. B., & Frisbie, D. A. (2001). *The Iowa Test of Basic Skills, Form A*. Itasca, IL: Riverside.

- Horn, J. L. (1982). The theory of fluid and crystallized intelligence in relation to concepts of cognitive psychology and aging in adulthood. In F. I. M. Craik & S. Trehub (Eds.), *Aging and cognitive processes* (pp. 237-278). New York, NY: Plenum.
- Hunsaker, S. L. (2000). Documenting gifted program results for key decision-makers. *Roeper Review*, 23(2), 80-83.
- Hunsaker, S. L., Abeel, L. B., & Callahan, C. M. (1991). *Instrument use in the identification of gifted and talented children*. Paper presented at the meeting of the Jacob K. Javits gifted and talented education program grant recipients, Washington, D.C.
- Jacob K. Javits Gifted and Talented Students Education Act (1988). P.L. 100-297, Title IV, Part B, Sec. 1101.
- Jenner, E., & Jenner, L. W. (2007). Results from a first-year evaluation of academic impacts of an after-school program for at-risk students. *Journal of Education for Students Placed at Risk*, 12(2), 213-237. doi: 10.1080/10824660701261144
- Jennings, J. F. (2000). Title I: Its legislative history and its promise. *The Phi Delta Kappan*, 81(7), 516-522.
- Johnsen, S. (1988, October). *The validity and reliability of instruments used in identifying gifted*

and talented students. Paper presented at the Conference of the National Association for Gifted Children. Orlando, FL.

Johnsen, S. (1997). Assessment beyond definitions. *Peabody Journal of Education*, 77(3/4), 136-142.

Johnsen, S. (Ed.). (2006). The college board [Special section]. *Gifted Child Today*, 29(4).

Johnsen, S. K. (2011). Evaluating the effectiveness of identification procedures. In S. Johnsen (Ed.), *Identifying gifted students: A practical guide* (2nd ed., pp. 151-159). Waco, TX: Prufrock Press.

Johnson, L. J., Karnes, M. B., & Carr, V. W. (1997). Providing services to children with gifts and disabilities: A critical need. In N. Colangelo & G.A. Davis (Eds.) *Handbook of gifted education* (2nd ed., pp 89-108). Boston, MA: Allyn and Bacon.

Jolly, J. (2004). *A conceptual history of gifted education: 1910-1940* (Unpublished doctoral dissertation). Baylor University, Waco, TX.

Jolly, J. L. (2008). A paradoxical point of view: Lewis M. Terman. *Gifted Child Today*, 31(2), 36-37.

Jolly, J. (2009). A resuscitation of gifted education. *American Educational History Journal*, 36(1), 37-52.

Jolly, J., & Bruno, J. (2010). Paul A. Witty: A friend of gifted children. *Gifted Child Today*, 33(4), 14-17.

Kane, H. D., & Brand, C. R. (2006). The variable importance of general intelligence (g) in the cognitive abilities of children and adolescents. *Educational Psychology*, 26(6), 751-767.

Kärkkäinen, R., Rätty, H., & Kasanen, K. (2009). Parents' perceptions of their child's resilience and competencies. *European Journal of Psychology of Education*, 24(3), 405-419.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence test* (2nd ed.). Circle Pines, MN: American Guidance Service.

Kerr, B. (1997). Developing talents in girls and young women. In N. Colangelo & G.A. Davis (Eds.), *Handbook of gifted education* (2nd ed., pp. 483-497). Boston, MA: Allyn and Bacon.

Kim, J. S., & White, T. G. (2008). Scaffolding voluntary summer reading for children in grades 3 to 5: An experimental study. *Scientific Studies of Reading*, 12(1), 1-23. doi: 10.1080/10888430701746849

Konstantopoulos, S., Modi, M., & Hedges, L. V. (2001). Who are America's gifted? *American Journal of Education*, 109(3), 344-382.

Kurdek, L. A., & Sinclair, R. J. (2001). Predicting reading and mathematics achievement in

fourth-grade children from kindergarten readiness scores. *Journal of Educational Psychology*, 93(3), 451-455.

Lee, S. Y., Matthews, M. S., & Olszewski-Kubilius, P. (2008). A national picture of talent search and talent search educational programs. *Gifted Child Quarterly*, 52(1), 55-69.

Leslie, M. (2000, July/August). The vexing legacy of Lewis Terman. *Stanford Magazine*.

Retrieved from <http://www.stanfordalumni.org/news/magazine/2000/julaug/articles/terman.html>

Lind, S. (2001). Overexcitability and the gifted. *The SENG Newsletter* (2001). Retrieved from http://www.sengifted.org/articles_social/Lind_OverexcitabilityAndTheGifted.shtml

Lohman, D. F. (2003). *The Woodcock-Johnson III and the Cognitive Abilities Test (Form 6): A concurrent validity study*. Retrieved from http://faculty.education.uiowa.edu/dlohman/pdf/CogAT_WJIII_final_2col%20r.pdf

Lohman, D. F. (2005a). An aptitude perspective on talent: Implications for identification of academically gifted minority students. *Journal for the Education for the Gifted*, 28(3), 333-360.

Lohman, D. F. (2005b). The role of nonverbal ability tests in identifying academically gifted students: An aptitude perspective. *Gifted Child Quarterly*, 49(2).

Lohman, D. F. (2006). Exploring perceptions and awareness of high ability.

Roepers Review, 29(1), 32-40.

Lohman, D. F., & Hagen, E. P. (2001). *Cognitive abilities test*. Itasca, IL: Riverside Publishing.

Lohman, D. F., Korb, K. A., & Lakin, J. M. (2008). Identifying academically gifted English-language learners using nonverbal tests: A comparison of the Raven, NNAT, and CogAT. *Gifted Child Quarterly*, 52(4), 275-296.

Louis, B., Subotnik, R. F., Breland, P. S., & Lewis, M. (2000). Establishing criteria for high ability versus selective admission to gifted programs: Implications for policy and practice. *Educational Psychology Review*, 12(3).

Loyd, B. H., Forsyth, R., & Hoover, H. D. (1981). Relationship of elementary and secondary school achievement test scores to later academic success. *Educational and Psychological Measurement*, 41, 1117-1124.

Lubinski, D., Benbow, C. P., Shea, D. L., Eftekhari-Sanjani, H., & Halvorson, M.B.J. (2001). Men and women at promise for scientific excellence: Similarity not dissimilarity. *Psychological Science*, 12(4), 309-317.

Luckasson, R., Bothwick-Duffy, S., Buntinx, W.H.E., Coulter, D. L., Craig, E. M., Reeve, A.,

- Schalock, R. L., Snell, M. E., Spitalnik, D. M., Spreat, S., & Tasse, M. J. (2002). *Mental retardation: Definition, classification, and systems of supports* (10th ed). Washington, DC: American Association on Mental Retardation.
- Lupkowski-Shoplik, A., & Assouline, S. (1994, February). Evidence of extreme mathematical precocity: Case studies of talented youths. *Roeper Review*, *16*(3), 144-151.
- Maddaus, G. F., & Rippey, R. M. (1966). Zeroing in on the STEP writing test: What does it tell a teacher? *Journal of Educational Measurement*, *3*(1), 19-25.
- Maehler, C., & Schuchardt, K. (2011). Working memory in children with learning disabilities: Rethinking the criterion of discrepancy. *International Journal of Disability, Development & Education*, *58*(1), 5-17. doi: 10.1080/1034912X.2011.547335
- Maker, C. J., & Nielson, A. B. (1995). *Teaching models in education of the gifted* (2nd ed). Austin, TX: Pro-Ed.
- Malley-Morrison, K., & Hines, D. A. (2007). Attending to the role of race/ethnicity in family violence research. *Journal of Interpersonal Violence*, *22*(8), 943-972.
- Markus, H. R., Steele, C. M., & Steele, D. M. (2000). Colorblindness as a barrier to inclusion: Assimilation and nonimmigrant minorities. *Daedalus*, *129*(4), 233-259.
- Marland, S. (1972). *Education of the gifted and talented: Report to the Congress of the United*

States by the U.S. Commissioner of Education. Washington, DC: Government Printing Office.

Mathews, F. N., & Burns, J. M. (1992). A parent evaluation of a public preschool gifted program. *Roeper Review*, 15(2), 69-73.

Matthews, M. S. (2006). Gifted students dropping out: Recent findings from a southeastern state. *Roeper Review*, 28(4), 216-223.

Matthews, M. S., & Shaunessy, E. (2010). Putting standards into practice: Evaluating the utility of the NAGC Pre-K-Grade 12 gifted program standards. *Gifted Child Quarterly*, 54(3), 159-167. doi: 10.1177/0016986209356708

Mattox, D. D. (1995). A predictive validity study of the Kindergarten Diagnostic Instrument. *The Ohio State University*, 1-166.

Mayr, E. (1982). *The growth of biological thought*. Cambridge, MA: Belknap Press.

McClellan, E. (1985). Defining giftedness. Retrieved from <http://eric.ed.gov/PDFS/ED262519.pdf>

McGonagill, B. (1997). Gifted education long-range planning: Using time wisely with TQM. *Roeper Review*, 19(4), 200.

McKenzie, R. G. (2009). Obscuring vital distinctions: The oversimplification of learning

- disabilities within RtI. *Learning Disability Quarterly*, 32(4), 203-215.
- Meservy, J. Z. (2008). The new classification of mental retardation as “Intellectual Disability”.
Brown University Child & Adolescent Behavior Letter, 24(11), 1-7.
- Milgram, R. (2000). Talent identification and development: An international
perspective. *Roepers Review*, 22(2), 76.
- Mills, C. J., & Barnett, L. B. (1992). The use of the Secondary School Admissions Test (SSAT)
to identify academically talented elementary school students. *Gifted Child Quarterly*, 36,
155-159.
- Mills, C. J., & Tissot, S. L. (1995). Identifying academic potential in students from under-
represented populations: Is using the Raven Progressive Matrices a good idea? *Gifted
Child Quarterly*, 39, 209-217.
- Morelock, M. J., & Feldman, D. H. (1997). High-IQ children, extreme precocity, and savant
syndrome. In N. Colangelo & G.A. Davis (Eds.) *Handbook of gifted education* (2nd ed.,
pp. 439-459). Boston, MA: Allyn and Bacon.
- Morisano, D., & Shore, B. M. (2010). Can personal goal setting tap the potential of the gifted
underachiever? *Roepers Review*, 32(4), 249-258. doi: 10.1080/02783193.2010.508156
- Naglieri, J. A. (1997). *NNAT multilevel technical manual*. San Antonio, TX: Psychological

Corporation.

Naglieri, J. A. (2003). *Naglieri Nonverbal Ability Test*. San Antonio, TX: Harcourt Brace & Company.

Nasca, D. (1988). *The use of non-verbal measures of intellectual functioning in identifying gifted children*. Retrieved from <http://eric.ed.gov/PDFS/ED296551.pdf>

National Association for Gifted Children. (2009). *State of the nation report 2008-2009*. Retrieved from <http://www.nagc.org/index.aspx?id=1051>

National Association for Gifted Children. (2010a). *Frequently asked questions*. Retrieved from <http://www.nagc.org/index.aspx?id=548>

National Association for Gifted Children. (2010b). *Pre-K-grade 12 gifted programming standards*. Retrieved from <http://www.nagc.org/index.aspx?id=546>

National Association for Gifted Children (2011). *State definitions of giftedness*. Retrieved from [http://nagc.org/uploadedFiles/Advocacy/State%20definitions%20\(8-24-10\).pdf](http://nagc.org/uploadedFiles/Advocacy/State%20definitions%20(8-24-10).pdf)

National Center for Learning Disabilities. (2011). *Learning disabilities definition*. Retrieved from <http://www.nclld.org/ld-basics>

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for*

educational reform. Washington, DC: U.S. Department of Education.

Neihart, M. (2006). Dimensions of underachievement, difficult contexts, and perceptions of self.

Roeper Review, 28(4), 196-202.

Neisser, U., Boodoo, G., Bouchard, Jr., T. J., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D.

F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51(2), 77-101.

Nelson, M. A., & Smith, S. W. (2001). External factors affecting gifted girls' academic and

career achievements. *Intervention in School and Clinic*, 37(1), 19-22. doi:

10.1177/105345120103700104

Nettelbeck, T., & Wilson, C. (2005). Intelligence and IQ: What teachers should know.

Educational Psychology, 25(6).

Neumeister, K.L.S., Williams, K. K., & Cross, T. L. (2009). Gifted high-school students'

perspectives on the development of perfectionism. *Roeper Review*, 31(4), 198-206. doi:

10.1080/0278319090317564

Newton, J. H., McIntosh, D. E., Dixon, F., Williams, T., & Youman, E. (2008). Assessing

giftedness in children: Comparing the accuracy of three shortened measures of intelligence to the Stanford-Binet Intelligence Scales, Fifth Edition. *Psychology in the Schools*, 45(6), 523-536. doi: 10.1002/pits.20321

Niederer, K., Irwin, R. J., Irwin, K. C. & Reilly, I. L. (2003). Identification of mathematically gifted children in New Zealand. *High Ability Studies*, 14(1).

Noble, K. D. (1987). The dilemma of the gifted woman. *Psychology of Women Quarterly*, 11(3), 367-378.

Noble, K. D., & Subotnik, R. F. (1999). To thine own self be true: A new model of female talent development. *Gifted Child Quarterly*, 43(3), 140-150.

Norton, B. (2006). Charles Spearman and the general factor in intelligence: Genesis and interpretation in the light of sociopersonal considerations. *Journal of the History of the Behavioral Sciences*, 15(2), 142-154. doi: 10.1002/1520-6696(197904)15:2<142::AID-JHBS2300150206>3.0.CO;2-X

Ntiri, D. W. (2009). Toward a functional and culturally salient definition of literacy. *Adult Basic Education and Literacy Journal*, 3(2), 97-104.

Omnibus Budget Reconciliation Act of 1981, Section 582, 42 USC 3842.

O'Neal, D., & Ringler, M. (2010). Broadening our view of linguistic diversity. *Phi Delta*

Kappan, 91(7), 48-52.

Otis, A. S., & Lennon, R. T. (1997). *Otis-Lennon School Ability Test* (7th ed.). San Antonio, TX: Harcourt Brace.

Parker, M. R., Jordan, K. R., Kirk, E. R., Aspiranti, K. B., & Bain, S. K. (2010). Publications in four gifted education journals from 2001 to 2006: An analysis of article types and authorship characteristics. *Roeper Review*, 32, 207-216.

Paulos, J. A. (2009). Do SAT scores really predict success? *ABC News*. Retrieved from <http://abcnews.go.com/print?id=98373>

Pestello, F. P. (1987). The social construction of grades. *Teaching Sociology*, 15(4), 414-417.

Peterson, J. S. (2002). A longitudinal study of post-high-school development in gifted individuals at risk for poor educational outcomes. *The Journal of Secondary Gifted Education*, 14(1), 6-18.

Peterson, J. S. (2006). Addressing counseling needs of gifted students. *Professional School Counseling*, 10(1), 43-51.

Peterson, J. S. (2009). Myth 17: Gifted and talented individuals do not have unique social and emotional needs. *Gifted Child Quarterly*, 53(4), 280-282.

Peterson, K.M.H. & Shinn, M. R. (2002). Severe discrepancy models: Which best explains

- school identification practices for learning disabilities? *School Psychology Review*, 31(4), 459-476.
- P.L. 507. 81st Congress. (1950). The National Science Foundation Act of 1950.
- Pyrczak, F., & Bruce, R. R. (2000). *Writing empirical research reports: A basic guide for students of the social and behavioral sciences* (3rd ed., pp 42-43). Los Angeles, CA: Pyrczak.
- Pyryt, M. (1996). IQ: Easy to bash, hard to replace. *Roeper Review*, 18(4), 255.
- Raborn, J. (2002). *Challenging schools' expectations of Native American students*. The National Research Center on the Gifted and Talented. Retrieved from <http://gifted.uconn.edu/nrcgt/newsletter/fall02/fall023.html>
- Raven, J. C., Court, J. H., & Raven, J. (1983). *Manual for Raven's Progressive Matrices and Vocabulary Scales, Section 4; Advanced Progressive Matrices, Sets I and II*. London, UK: H.K. Lewis.
- Rebhorn, L.S., & Miles, D. D. (1999). High-stakes testing: Barrier to gifted girls in mathematics and science? *School Science and Mathematics*, 99.
- Regulations for the Education for Economic Security Act of 1984. Part 208, Section 208.3.
- Reis, S. M., (2003). Gifted girls, twenty-five years later: Hopes realized and new challenges

- found. *Roeper Review*, 25(4), 154-159.
- Reis, S. M., & McCoach, D. (2000). The underachievement of gifted students: What do we know and where do we go? *Gifted Child Quarterly*, 44(3), 152-170.
- Reis, S. M. & Renzulli, J. S. (2010). Is there still a need for gifted education? An examination of current research. *Learning & Individual Differences*, 20(4), 308-317.
- Renzulli, J. S. (1986). The three-ring conception of giftedness: A developmental model for creative productivity. In R. J. Sternberg & J. Davidson (Eds.), *Conceptions of giftedness*. (pp. 53-92). New York, NY: Cambridge University Press.
- Renzulli, J. S. (2002). Emerging conceptions of giftedness: Building a bridge to the new century. *Exceptionality*, 10(2), 67-75.
- Renzulli, J. S., & Park, S. (2000). Gifted dropouts: The who and the why. *Gifted Child Quarterly*, 44, 261-271.
- Rimm, S. B. (1996) Parenting for achievement. *Roeper Review*, 19(1), 57-59.
- Rimm, S. B. (1997a). An underachievement epidemic. *Educational Leadership*, 54(7), 18-22.
- Rimm, S. B. (1997b). Underachievement syndrome: A national epidemic. In N. Colangelo & G.A. Davis (Eds.). *Handbook of gifted education* (2nd ed., pp. 416-434). Boston, MA: Allyn and Bacon.

- Rimm, S. B. (2012). Bright kids, poor grades: And what you can do about it. Retrieved from http://www.sylviarimm.com/article_brightkids.html
- Robins, J., & Jolly, J. (2011). Technical information regarding assessment. In S. Johnsen (Ed.) *Identifying gifted students: A practical guide* (2nd ed., pp 75-118). Waco, TX: Prufrock Press.
- Robbins, R. (1991). American Indian gifted and talented students: Their problems and proposed solutions. *Journal of American Indian Education, 31*(1). Retrieved from <http://jaie.asu.edu/v31/V31S1Ame.htm>
- Roberts, J. L. (1999). The top 10 events creating gifted education for the new century. *Gifted Child Today, 22*(6), 53-55.
- Robinson, N. M., Zigler, E., & Gallagher, J. J. (2000). Two tails of the normal curve: Similarities and differences in the study of mental retardation and giftedness. *American Psychologist, 55*, 1413–1424. doi:10.1037/0003-066X.55.12.1413
- Robinson, R. W., & Miller, D. C. (1990). *Kindergarten Diagnostic Instrument: Administration manual*. Denton, TX: KIDs, Inc.
- Roeper, A. (2003). The young gifted girl: A contemporary view. *Roeper Review, 25*.
- Roid, G. H. (2003). *Stanford-Binet Intelligence Scales* (5th ed.). Itasca, IL: Riverside

Ryser, G. R. (2011). Fairness in testing and nonbiased assessment. In S. Johnsen (Ed.).

Identifying gifted students: A practical guide (2nd ed., pp. 63-74). Waco, TX: Prufrock Press.

Sadker, D. (2002). An educator's primer on the gender war. *The Phi Delta Kappan*, 84(3), 235-244.

Samuels, C. A. (2007). States seen renewing focus on education of gifted. *Education Week*, 26(32), 20-23.

Santelices, M. V., & Wilson, M. (2010). Unfair treatment? The case of Freedle, the SAT, and the standardization approach to differential item functioning. *Harvard Educational Review*, 80(1), 106-134.

Scherbaum, C. A., & Goldstein, H. W. (2008). Examining the relationship between race-based differential item functioning and item difficulty. *Educational & Psychological Measurement*, 68(4), 537-553.

Schiever, S. W., & Maker, C. J. (1997). Enrichment and acceleration: An overview and new directions. In N. Colangelo & G.A. Davis (Eds.). *Handbook of gifted education* (2nd ed., pp. 89-108). Boston, MA: Allyn and Bacon.

Schilling, S. G., Carlisle, J. F., Scott, S. E., & Ji Zeng. (2007). Are fluency measures accurate

- predictors of reading achievement? *Elementary School Journal*, 107(5), 429-448.
- Schlachter, S. (2008). Diagnosis, treatment, and educational implications for students with attention-deficit/hyperactivity disorder in the United States, Australia, and the United Kingdom. *Peabody Journal of Education*, 83(1), 154-169.
- Schober, B., Reimann, R., & Wagner, P. (2004). Is research on gender-specific underachievement in gifted girls an obsolete topic? New findings on an often discussed issue. *High Ability Studies*, 15(1), 43-62.
- Siegle, D. (2012). *Parenting strategies to motivate underachieving gifted students*. Retrieved from <http://www.tip.duke.edu/node/817>
- Siegle, D., & Reis, S. M. (1998). Gender differences in teacher and student perceptions of gifted students' ability and effort. *Gifted Child Quarterly*, 42, 39-47.
- Silverman, L. K. (1997). Family counseling with the gifted. In N. Colangelo & G.A. Davis (Eds.). *Handbook of gifted education* (2nd ed., pp. 382-397). Boston, MA: Allyn and Bacon.
- Sisk, D. (1990). The state of gifted education. *Music Educators Journal*, 76(7), 35.
- Smith, P. (2001) Mentors as gate-keepers: An exploration of professional formation. *Educational Review*, 53(3), 313-325.

- Spearman, C. (1925). Some issues in the theory of “g” (including the Law of Diminishing Returns) 1. *Nature*, *116*, 436-439. doi: 10.1038/116436a0
- Spencer, H. E. (1996). Mathematical SAT test scores and college chemistry grades. *Journal of Chemical Education*, *73*(12), 1150.
- Spencer, M. B., Dupree, D., Swanson, D. P., & Cunningham, M. (1998). The influence of physical maturation and hassles on African American adolescents’ learning behaviors. *Journal of Comparative Family Studies*, *29*(1), 189-200.
- Stanley, G., & Baines, L. (2004). No more shopping for grades at B-Mart: Re-establishing grades as indicators of academic performance. *The Clearing House*, *77*(3), 101-104.
- Stanley, J. C. (2005). A quiet revolution: Finding boys and girls who reason exceptionally well and/or verbally and help them get the supplemental educational opportunities they need. *High Ability Studies*, *16*(1), 5-14. doi: 10.1080/13598130500115114
- Stanovich, K. E. (2005). The future of a mistake: Will discrepancy measurement continue to make the learning disabilities field a pseudoscience? *Learning Disability Quarterly*, *28*(2), 103-105.
- Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York, NY: Cambridge University Press.

Sternberg, R. J. (1996). Myths, countermyths, and truths about intelligence. *Educational Researcher*, 25(2), 11-16.

Sternberg, R. J. (1997) A triarchic view of giftedness: Theory and practice. In N. Colangelo & G.A. Davis (Eds.). *Handbook of gifted education* (2nd ed., pp. 89-108). Boston, MA: Allyn and Bacon.

Sternberg, R. J. (2008). Applying psychological theories to educational practice. *American Educational Research Journal*, 45(1), 150-165.

Sternberg, R. J., & Clinkenbeard, P. (1995, May). The triarchic model applied to identifying, teaching, and assessing gifted children. *Roeper Review*, 17(4), 255.

Stewart, E. D. (1999). An American century of roots and signposts in gifted and talented education. *Gifted Child Today*, 22(6), 56-57.

Stringer, N. (2008). Aptitude tests versus school exams as selection tools for higher education and the case for assessing educational achievement in context. *Research Papers in Education*, 23(1), 53-68.

Strip, C. A., & Hirsch, G. (2000). *Helping gifted children soar: A practical guide for parents and teachers*. Scottsdale, AZ: Great Potential Press.

Subotnik, R. F., & Arnold, K. D. (1995). Gifted is as gifted does but what does gifted do? *Roeper*

Review, 18(1) 4-6.

Subotnik, R. F., Olszewski-Kubilius, P., & Worrell, F. C. (2011). Rethinking giftedness and gifted education: A proposed direction forward based on psychological science. *Psychological Science in the Public Interest, 12(1)*, 3-54. doi: 10.1177/1529100611418056

Swiatek, M. A., & Lupkowski-Shoplik, A. (2000). Gender differences in academic attitudes among gifted elementary school students. *Journal for the Education of the Gifted, 23*, 360-377.

Tang, M., & Neber, H. (2008). Motivation and self-regulated science learning in high-achieving students: Differences related to nation, gender, and grade-level. *High Ability Studies, 19(2)*, 103-116. doi: 10.1080/13598130802503959

Taylor, H. E., & Larson, S. (1998). Teaching children with ADHD: What do elementary and middle school social studies teachers need to know? *Social Studies, 89(4)*, 161-164.

Terman, L. M. (1920). The use of intelligence tests in the grading of school children. *Journal of Educational Research, 1(1)*, 20-32.

Terman, L. M. (1921). A symposium: Intelligence and its measurement. *Journal of Educational Psychology, 12*, 127-133.

Terman, L. M. (1925). *Genetic studies of genius* (Vol. 1). Stanford, CA: Stanford University Press.

Terman, L. M. (1939). Educational suggestions from follow-up studies of intellectually gifted children. *Journal of Educational Sociology*, 13(2), 82-89.

Texas Education Agency (2010). Public education information management systems: *At-risk indicator code* (E0919). Retrieved from <http://ritter.tea.state.tx.us/peims/standards/wedspre/index.html?e0919>

Texas Education Code Ch. §29, 29.121. Education Programs for Gifted and Talented Students. 2009 Texas Educational Programs 74th Legislature. 1 September 1995.

Thorndike, R. L. (1975). Mr. Binet's test 70 years later. *Educational Researcher*, 4(5), 3-7.

Thurstone, L. L. (1946). Theories of intelligence. *The Scientific Monthly*, 62(2), 101-112.

Tomlinson, C. A. (2003). Differentiating instruction for academic diversity. In J. M. Cooper (Ed.), *Classroom teaching skills* (7th ed., pp. 149-180). Boston, MA: Houghton Mifflin.

Tomlinson, C. A., & Callahan, C. M. (1994). Planning effective evaluations for programs for the gifted. *Roeper Review*, 17(1), 46-52.

Tyson, K., Darity, Jr., W., & Castellino, D. R. (2005). It's not "a black thing": Understanding the

- burden of acting white and other dilemmas of high achievement. *American Sociological Review*, 70(4), 582-605.
- U.S. Department of Education (1977). Assistance to states for education for handicapped children: Procedures for evaluating specific learning disabilities. *Federal Register*, 42, 62082-62085.
- U.S. Department of Education. (1993). *National excellence: A case for developing america's talent*. Retrieved from <http://www.ed.gov/pubs/DevTalent/part2.html>
- U.S. Department of Education. (2002). *No Child Left Behind Act of 2001* (159, 115 Stat PUBLIC LAW 107-110). Washington, DC: U.S. Department of Education.
- Van Tassel-Baska, J. (1997). What matters in curriculum for gifted learners: Reflections on theory, research, and practice. In N. Colangelo & G.A. Davis (Eds.). *Handbook of gifted education* (2nd ed., pp. 126-135). Boston, MA: Allyn and Bacon.
- Van Tassel-Baska, J. (2001). Assessment of gifted student learning in the language arts. *The Journal of Secondary Gifted Education*, 13(2), 67-72.
- Van Tassel-Baska, J. (2006). A content analysis of evaluation findings across 20 gifted programs: A clarion call for enhanced gifted program development. *Gifted Child Quarterly*, 50(3), 199-210.

- Vernon, P. E. (1940). *The measurement of abilities*. London: University of London Press.
- Wechsler, D. (2003). *Wechsler Intelligence Scale for Children – Fourth Edition*. San Antonio, TX: Psychological Corporation.
- Weinberg, R.A. (1989). Intelligence and IQ: Landmark issues and great debates. *American Psychologist*, 44(2), 98-104.
- Williams, R. H., Zimmerman, D. W., Zumbo, B. D., & Ross, D. (2003). Charles Spearman: British behavioral scientist. *Human Nature Review*, 3, 114-118.
- Willingham, W.W., Pollack, J.M., & Lewis, C. (2002). Grades and test scores: Accounting for observed differences. *Journal of Educational Measurement*, 39(1), 1-37.
- Wilson, J. L. (2007). Virtual teaming: Placing preservice middle level teachers on interdisciplinary teams. *Research in Middle Level Education Online*, 31(3). Retrieved from <http://www.nmsa.org/Publications/RMLEOnline/Articles/Vol31No3/tabid/1559/Default.aspx>
- Witty, P. A. (1962). The gifted and the creative pupil. *Education*, 82, 451–459.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson Test of Cognitive Ability, Third Edition*. Itasca, IL: Riverside Publishing.
- Yssel, N. Prater, M., & Smith, D. (2010). How can such a smart kid not get it? *Gifted Child*

Today, 33(1), 54-61.

Appendices

Appendix A



Gifted/Talented Program Middle School English Identification Matrix 6-8

Student #: _____ Birth Date: _____
 Current Grade Level _____ Elementary School: _____
 Testing for Grade: _____ Middle School: _____
 Home Address: _____ Zip Code: _____
 Home Phone: _____ Work Phone: _____

Quantitative Data					N/A
Academic Aptitude (Percentile) Title: _____ V- Q- Date: _____ NV-	99-98	97-95	94-92	91-89	
Nonverbal Ability Test (Percentile) Title: _____ Date: _____	99-96	97-90	89-85	84-80	
Reading Achievement Title: _____ (Percentile) Date: _____	99-98	97-95	94-92	91-89	
STEP-Writing (Percentile) Date: _____	99-98	97-95	94-92	91-89	
Commended TAKS Reading (5 th , 6 th , 7 th , 8 th)	YES		NO		No Data
Commended TAKS Writing (4 th , 7 th)	YES		NO		No Data
Qualitative Data					
English Grade Average (Percent)	98+	97-95	94-92	91-89	
Behavioral Characteristics Evaluation (Percent)	100-94	93-86	85-76	75-66	
Additional Anecdotal Evidence Attached	Yes		No		

_____ Qualifies Does Not Qualify for G/T English.

SELECTION COMMITTEE:

 _____ Date: _____

Appendix B



Gifted/Talented

TEACHER CHECKLIST CHARACTERISTICS OF G/T STUDENTS SECONDARY ENGLISH

Student: _____	School: _____	Grade Level: _____
Person completing form: _____		Date: _____

Directions: Write the number in the column that best represents your observation of the following characteristics in the above named student. Please total the numbers in each column and write the final total on the line provided.

Characteristics	Examples	Superior	Above Average	Average	Weak	Does Not Exist
		4	3	2	1	0
1. Grasps and retains knowledge	<input type="checkbox"/> comprehends meanings <input type="checkbox"/> questions critically <input type="checkbox"/> transfers learning to new situations					
2. Conveys ideas effectively	<input type="checkbox"/> follows logical sequence and order <input type="checkbox"/> has extensive vocabulary and uses it appropriately					
3. Shows skill in abstract thinking	<input type="checkbox"/> makes generalizations <input type="checkbox"/> senses cause and effect <input type="checkbox"/> recognizes relationships <input type="checkbox"/> understands and applies rules <input type="checkbox"/> foresees new possibilities					
4. Has creative and inventive power	<input type="checkbox"/> shows curiosity and originality <input type="checkbox"/> enjoys experimenting & finding ways to extend ideas <input type="checkbox"/> uses trial and error method					
5. Exhibits power to work independently	<input type="checkbox"/> shows ability to plan, organize, execute, & judge <input type="checkbox"/> show self-reliance when meeting challenges <input type="checkbox"/> knows when, where, and how to seek help					
6. Adjusts easily to new situations	<input type="checkbox"/> understands and accepts reasons for change <input type="checkbox"/> anticipates outcomes <input type="checkbox"/> maintains optimistic attitude toward new adventures <input type="checkbox"/> accepts challenges					
7. Task commitment	<input type="checkbox"/> shows strong self-motivation <input type="checkbox"/> uses class time wisely & completes work <input type="checkbox"/> shows perseverance & a will to succeed					
TOTALS:						

Total Points: _____

	YES	NO
This student demonstrates the ability to work beyond his/her current peers.		

*Teachers use the space below to write a brief summary of why this student should be considered for G/T.

Appendix C

University Committee for the Protection
of Human Subjects in Research
University of New Orleans

Campus Correspondence

Principal Investigator: Mary Cronin, Paul Bole

Co-Investigator: Colleen D. Lindsey

Date: September 22, 2011

Protocol Title: "An Investigation of the Combined Assessments Used As
Entrance Criteria for a Gifted English Middle School Program"

IRB#: 04Sep11

The IRB has deemed that the research and procedures described in this protocol application are exempt from federal regulations under 45 CFR 46.101 category 1. This minimal-risk study will be conducted in established or commonly accepted educational settings, involving normal educational practices and will entail research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.

Exempt protocols do not have an expiration date; however, if there are any changes made to this protocol that may cause it to be no longer exempt from CFR 46, the IRB requires another standard application from the investigator(s) which should provide the same information that is in this application with changes that may have changed the exempt status.

If an adverse, unforeseen event occurs (e.g., physical, social, or emotional harm), you are required to inform the IRB as soon as possible after the event.

Best wishes on your project.

Sincerely,

Robert D. Laird, Ph.D., Chair

Committee for the Protection of Human Subjects in Research

Appendix D

November 3, 2011

Colleen Lindsey
Bradley Middle School
14819 Heimer Road
San Antonio, TX 78232

Dear Ms. Lindsey:

The Department of Research and Information Technologies has approved your request to use Gifted/Talented data from Bradley Middle School for the years of 2006-2009 under the guidance of G/T Coordinator, Kay Stotts. Please note the additional terms and conditions of this approval:

- 1. The study makes minimal interruptions to the regular school program and makes no undue demands upon the time of students, teachers, administrators, or other district personnel.**
- 2. The district reserves the right to decline future solicitations for this project.**
- 3. No student names are used in any reports or publications.**
- 4. Student identity is removed before analysis and destroyed at the end of the study.**

Please note – participation in your study by individual district personnel is strictly voluntary. A copy of this letter should accompany your solicitation to district personnel. Should you have questions or need additional information, please contact me at swatso8@neisd.net or by telephone at (210) 407-0559.

Sincerely,

Susan Watson, Ph.D.
Research Analyst

Cc: Mark Scheffler, Associate Superintendent, Division of Campus Support and Human Res.
Janna Hawkins, Associate Superintendent, Division of Instruction & Tech. Services
Mike Lara, Executive Director for Research & Information Technologies
Jeff Kurth, Director, Research & Information Technologies
Justin Oxley, Principal, Bradley Middle School
Kay Stotts, G/T Coordinator K-12

Appendix E

Date: November 3, 2011

Memo To: Colleen Lindsey, Bradley Middle School

From: Kay Stotts, G/T Coordinator K-12

Through: Janna Hawkins, Associate Superintendent Division of Instruction & Technology Services

Subject: Permission for Research Project for Doctorial Study

I will provide the matrices of the Bradley students who tested and qualified for G/T English during the '07-'08, '08-'09, '09-'10 school years. They may be viewed in my office with me present during the viewing.

It is my understanding no student names will be collected with the information that is gathered during the viewing.

I also understand NEISD Superintendent Dr. Brian Gotardy and Bradley Principal, Justin Oxley have granted permission for this research project.

Please, contact me to arrange a date for the viewing. Kay Stotts 407-0327

Appendix F

November 10, 2011

To Whom It May Concern:

I grant permission for Colleen Lindsey to conduct a research project using data from Bradley gifted English students during the 2007-08, 2008-09, and 2009-2010 school years.

It is my understanding this study will remove student names when the data is collected, so that students' identities will be protected. Students' identity will be further protected when all data is destroyed at the conclusion of the study.

I also understand that NEISD Superintendent Dr. Brian Gottardy, through the Department of Research and Information Technologies, and Kay Stotts, Director of Gifted Education, have granted permission for this research project.

Sincerely,

Justin Oxley

Principal, Bradley Middle School

Vita

Colleen Lindsey is an educator of gifted children. She obtained her bachelor's degree in elementary education, with an emphasis in reading from Baylor University in 1991. She obtained her master's degree in curriculum and instruction, with an emphasis in adolescence and reading from Baylor University in 1992. She joined the University of New Orleans special education graduate program to pursue a Ph.D. in gifted education in 1998. She took a hiatus from the program for several years, but returned in 2009 to complete her degree. She teaches in San Antonio, Texas.