The Relationship Between ACT Scores and Core GPAs for South-Central Kansas High School Graduates in the Class of 2017

Loren B. Scarbrough
B.A., Calvin College, 2009
M.A., Wichita State University, 2013

Submitted to the Graduate Department and Faculty of the School of Education of Baker University in partial fulfillment of the requirements for the degree of Doctor of Education in Educational Leadership

Sharon Zoellner, Ph.D.
Major Advisor

Alison Banikowski, Ph.D.

Allen Jantz, Ed.D.

Date Defended: June 26, 2018

Abstract

The purpose of this study was to examine the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs calculated from all English, mathematics, history, and science courses in grades 9 through 12 in high schools in three south-central Kansas school districts. An additional purpose of the study was to explore how gender, race, socioeconomic status, and high school community type affected the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs for students in grades 9 through 12 in three south-central Kansas districts. A quantitative correlational research design was used in this study, which utilized archived student data from three south-central Kansas school districts to analyze the relationship between ACT scores and core GPAs.

The results of the data analysis indicated that students' ACT scores and core GPAs have a moderately strong correlation. Student gender did not affect the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs. However, race affected the relationship between English and mathematics ACT scores and core GPAs. The relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs was significantly affected by student socioeconomic status. Furthermore, community type significantly affected the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs. The findings from this study should encourage district and building leaders and teachers to scrutinize grading practices and examine the discrepancy in correlation between ACT scores and core GPAs based on socioeconomic status and community type.

Dedication

This dissertation is dedicated to my grandfather, Gerald Taylor, who always said, "Loren Brooke, you keep going to school now. Promise me you will." in his perfect Mississippi twang. Well, Daddy Gerald, I kept my promise and fulfilled our dream. While I wish you could be here to see it, I know you are beaming with pride in heaven. Hail State!

To my students and athletes, may this dissertation represent the possibility of perfectly matched dreams, perseverance, and grit. I have always maintained that being a role model is the most powerful form of educating and I hope that while pursuing my dreams, I was able to encourage yours. You have all made me a better educator, and the memories we made within the walls of Southeast High School and on the softball diamond will not be forgotten. Together We Did.

Acknowledgements

While the physical degree will have my name on it, writing a dissertation and completing a doctoral program is far from a solitary accomplishment. I would be remiss if I did not take the time to acknowledge individual contributions to my growth as an educator over the last three years.

To my mom, Tina Fawver, who excitedly volunteered to edit drafts and consistently cheered every major and minor accomplishment along this journey, thank you for being my biggest fan and constant role model.

Dusty, despite my selfishness with free time you showed continual understanding of my commitment to seeing this process through. There's no one else I would rather go through this journey with. I love you.

To Kinsey, for always bringing the sunshine. Thank you for your gift of friendship and laughter during this grueling process. You consistently listened to my frustrations and responded with joy; that was and is such an invaluable gift.

To my mentor, Lori Doyle, thank you for being honest with me about my strengths and areas of growth. The support you have shown me throughout this process has been nothing short of extraordinary. I am proud to have learned from your guidance and wisdom. Do not be surprised when your phone rings years from now and it is me still seeking your advice!

Thank you to my advisor, research analyst, and defense committee: Dr. Sharon Zoellner, Dr. Peg Waterman, Dr. Alison Banikowski, and Dr. Allen Jantz. You spent countless hours thoroughly editing my work and providing valuable feedback. Your patience and professionalism throughout this process consistently demonstrated your

investment in my academic success. Each of you is a standard-bearer for educational excellence.

To Christine Slane, Don Bonilla, and Rob Maddux, thank you for your understanding and support when I left practice early every Wednesday to go to class and work on this dissertation. Your ability to sustain a culture of excellence simultaneously buoyed the success of our athletic programs and my ability to complete this lifelong dream.

Dr. Lisa Lutz and Dr. John Black, thank you for supervising my directed field experiences. Despite hectic schedules of your own, you took the time to model the demeanor, planning, and professionalism required at the highest positions in PreK-12 education. What I learned from both of you will certainly carry my career forward.

Finally, to Ron Barry and Anthony Fulton, thank you for being my partners in this process and for attempting to teach me that celebratory dance I am still working on mastering. Baker University PreK-12 Educational Leadership Cohort 16, thank you for challenging me personally and professionally over the last three years. It is an honor to call each of you a colleague and a friend; one clap!

Table of Contents

List of Tables

Chapter 1

Introduction

By the late 19th century, Yale University, Harvard University, and Mount Holyoke College had developed the percentage-based grading system for classifying student academic performance still in use in the 21st century (Zirkel, 2007). When compulsory attendance laws were enacted in the early 20th century, elementary and secondary institutions turned to the grading paradigm developed by Yale, Harvard, and Mount Holyoke as a mechanism for placing the swelling student population in public education into academically similar groups (Kirschenbaum, Simon, & Napier, 1971). The Yale, Harvard, and Holyoke grading paradigm emphasizes points earned over standards mastered and the amount of work a student completes has become synonymous with rigor (Kohn, 2002). Since 1989, Canady and Hotchkiss have maintained that grading paradigms and educational expectations are not correlated with the demands a global society will place on students since basic declarative knowledge holds more value than procedural fluency and problem-solving. Meanwhile, students' future jobs will require critical thinking, complex tasks, sophisticated collaboration, and refined communication skills (Canady & Hotchkiss, 1989).

Despite the longevity of the grading paradigm, there have been questions about the legitimacy of grades for decades as teachers have worried that students have mistakenly come to understand that effort, not scholarship, is what earns a grade of A (Varlas, 2013). "In the present practice Grades A and B are sometimes given too readily – Grade A for work of not high merit, and Grade B for work not far above mediocrity" (Lewis, 2006, p. 115). Unlike Lewis' (2006) and Varlas' (2013) claims about poor grading practices, Goodwin (2011) argued “the fact that so many people could worry about [grading practices] makes one wonder whether concerns are grounded in reality or are merely generational grumblings about the declining standards of youth” (p. 80).

As a result of the uncertainty surrounding grades, the ACT, composed of four tests (English, mathematics, reading, and science), was designed to measure a student’s capacity to succeed in rigorous post-secondary courses and to determine the academic skills a student developed via secondary coursework (ACT, 2005). Composite ACT scores, the average of the four tests for each student, have been used by post-secondary institutions to determine course placement and college admissions along with grade point average (GPA) and other application components (ACT, 2005). The ACT Board of Directors, who regulates the ACT, uses statistical adjustments through the National Curriculum Survey to ensure all versions of the test are comparable in content and difficulty over time (ACT, 2005). Since the ACT is “constructed to measure the same content from year to year, and technical procedures are employed to ensure it has the same statistical properties over time” (ACT, 2005, p. 1), researchers have compared composite ACT scores to high school GPAs to respond to claims about the validity of grades by determining if the relationship has gotten stronger or weaker over time. However, school district officials cannot use the comparison between composite ACT scores and core high school GPA as an indicator of grade validity or invalidity, without first establishing how the two variables are related within their specific district.

Background

In 2004, Woodruff and Ziomek compared ACT scores to course GPAs for high school seniors who tested between 1991 and 2003. The 12-year period saw a 0.25

increase in English GPAs on a 4-point scale (6.25%) with a 0.22 increase on ACT English scores on a 36-point scale (0.61%) (Woodruff & Ziomek, 2004). Comparably, mathematics GPAs rose 0.24 points (6.00%), while ACT mathematics scores increased by 0.51 (1.42%) (Woodruff & Ziomek, 2004). Since ACT scores and high school GPA change rates were dissimilar, but ACT test difficulty remained consistent, ACT (2005) concluded poor grading practices, not achievement, accounted for the discrepancy. Three years later, Schmidt (2007) reached a similar conclusion when he determined that despite a decline in student performance on the National Assessment of Educational Progress (NAEP) every year between 1992 and 2005, student GPAs rose from 2.68 to 2.98. Zhang and Sanchez (2013) followed Woodruff and Ziomek's (2004) study comparing high school GPAs and ACT scores with differing results for students graduating after 2003. "From 1991 to 2003 there was an increase in [core] high school GPA for the selected ACT scores. After 2003, there was comparatively little change in the [core] high school GPA for the same scores" (Zhang & Sanchez, 2013, p. 10).

Three school districts from south-central Kansas agreed to participate in this study: Sunflower Public Schools, Daisy Public Schools, and Marigold Public Schools. Sunflower Public Schools is an urban district that had an enrollment of 50,566 students or 11% of all publicly educated students in the state of Kansas in 2017 (Sunflower Public Schools, 2017). Daisy Public Schools is a suburban district that had an enrollment of 7,173 students in eleven school buildings in 2017 (Daisy Public Schools, 2017). Marigold Public Schools is a rural district that served 2,295 students in 2017 (Marigold Public Schools, 2017). Daisy Public School graduates consistently perform above the Kansas state average composite ACT score, while Marigold Public School students average the same as the entire state and Sunflower Public School students average a composite ACT score below the Kansas state average. Table 1 presents the average ACT score of the participating districts compared to the average ACT score for the state of Kansas from 2013 to 2017.

Table 1

Districts and State Average Composite ACT Scores 2013-2017

School Year	Sunflower	Daisy	Marigold	Kansas
2013	19.8	22.7	22.3	21.9
2014	19.7	22.9	21.6	22.0
2015	19.6	22.5	21.6	21.9
2016	19.7	22.0	21.8	21.8
2017	19.3	21.2	21.4	21.7

Note. Adapted from *KSDE Report Card.* by KSDE, 2017. Retrieved from http://ksreportcard.ksde.org/act_scores.aspx?org_no=D0[redacted]&rptType=2, http://ksreportcard.ksde.org/act_scores.aspx?org_no=D0[redacted]&rptType=2, and http://ksreportcard.ksde.org/act_scores.aspx?org_no=D0[redacted]&rptType=2

Sunflower Public Schools, Daisy Public Schools, and Marigold Public Schools do not require all students to take the ACT, so the scores indicated in the table are the averages for students who consider themselves college bound, as self-reported on the ACT. The average ACT score in the state of Kansas is 21.7, which indicates the typical Kansas high school graduate would earn a B or higher in an entry level college course in English, mathematics, literature, or science (KSDE, 2017a, 2017b, 2017c). However, there is a large discrepancy in districts across Kansas (KSDE, 2017a, 2017b, 2017c). In 2015, McDaniel noted that over 27% of high school graduates were required to enroll in remedial post-secondary coursework, despite an average ACT score which indicates the typical Kansas high school graduate is college ready.

Statement of the Problem

The landscape of American education has changed drastically since the 19th century, but the original grading paradigm is still prevalent in today's educational practices (Vatterott, 2015). Grading has been at the heart of the kindergarten through post-secondary experience for 120 years. Regardless of the original validity of criteria used, grades have come to epitomize academic ability, which leaves teachers, students, parents, and college admissions staff to question whether a grade is truly indicative of mastery of standards (Pope, 2011). While extensive research has been conducted on the predictive ability of ACT scores and high school GPAs for college readiness and college success (Beecher & Fisher, 1999; Caruth & Caruth, 2013; Coyle & Pillow, 2008; Geiser & Santelices, 2007; Sawyer, 2013; Stumpf & Stanley, 2002; Westrick, Le, Robbins, Radunzel, & Schmidt, 2015; Zellner, 2008), there is less research on the relationship between composite ACT scores and high school GPAs (Woodruff & Ziomek, 2004; Zhang & Sanchez, 2013). Despite not establishing the nature of the relationship between ACT scores and core high school GPA, some school districts and researchers use ACT scores as the baseline for their studies of grading practices. However before school districts could use ACT scores to measure the accuracy of grading practices in core content classes (English, mathematics, history, and science), the districts needed to know whether there was any relationship between ACT scores and core GPAs.

Purpose of the Study

This study was conducted to examine the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs calculated from all English, mathematics, history, and science courses in grades 9 through 12 in high schools in three south-central Kansas school districts. An additional purpose of the study was to explore how gender, race, socioeconomic status, and high school community type affected the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs for students in grades 9 through 12 in three south-central Kansas districts.

Significance of the Study

Similar to Geiser and Santelices' (2007) study on the predictive ability of high school grades for college success, the results of this study may foster a review of effective grading practices, encourage the expansion of professional development models related to grading practices, and promote conversation regarding the validity of using GPAs and standardized test scores to measure grading practices if the measures are not as strongly correlated as educators assume. Additionally, this study may allow schools to identify demographic indicators of standardized test failure and respond more quickly to student needs that are masked by GPAs, since it is progressively more important for principals and superintendents to evaluate the relationship between student grades and standardized test scores for their schools and districts to better support accurate assessment of the mastery of core content standards. Woodruff and Ziomek (2004) and Zhang and Sanchez (2013) published two prominent studies on the relationship between composite ACT scores and GPAs, but the studies had differing results for test scores after 2003. The results of the current study contribute to the overall body of knowledge where research results have not been congruent.

Delimitations

According to Lunenburg and Irby (2008), "delimitations are self-imposed boundaries set by the researcher on the purpose and scope of the study" (p. 134). The following delimitations were imposed on this study:

1. The sample data for this study included ACT scores of students who met graduation requirements as outlined by the Kansas State Department of Education (KSDE, 2015) and graduated from three school districts in south-central Kansas during the 2016-2017 school year.
2. The data from this study included core GPAs calculated from the core courses of English, mathematics, history, and science of students who met the first delimitation.

Assumptions

"Assumptions are postulates, premises, and propositions that are accepted as operational for purposes of the research" (Lunenburg & Irby, 2008, p. 135). Additionally, the assumptions lend support to the recommendations by providing meaning to the conclusions drawn in the study (Lunenburg & Irby, 2008). The following assumptions were made during this study:

1. All participating high schools provided data that accurately reflected students' ACT scores, core GPAs (English, mathematics, history, and science) in grades 9 through 12, and community type.
2. The districts identified and coded student gender, race, and socioeconomic status with fidelity.

3. The sample selected for this study was representative of students who met graduation requirements as outlined by the KSDE (2015) in high schools in south-central Kansas.

Research Questions

According to Lunenburg and Irby (2008), "research questions…are critical components of the dissertation" (p. 126). Research questions identify the crux of the study and the questions that remain unanswered in the literature (Lunenburg & Irby, 2008). The following research questions guided this study:

RQ1. To what extent is there a relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs for students in grades 9 through 12 in three south-central Kansas districts?

RQ2. To what extent is the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs for students in grades 9 through 12 in three south-central Kansas districts affected by gender, race, socioeconomic status, and community type?

Definition of Terms

"Key terms need to be clarified if they are paramount to the study and referenced or used continuously throughout the dissertation" (Lunenburg & Irby, 2008, p. 118). The following terms are defined for this study:

American College Test (ACT). The ACT is a standardized college admissions test developed by ACT to measure student understanding of English, mathematics, reading, and science (ACT, 2014).

Community type. The designation of urban, suburban, or rural was determined by geographic location and Education Demographic and Geographic Estimates (EDGE) program categorization of the school district from which a student graduated (Geverdt, 2015).

Core grade point average (core GPA). Garton, Ball, and Dyer (2002) defined core GPA as a non-weighted numeric average of all earned grades in the four major academic areas during grades 9 through 12: English, mathematics, history, and science. Students receive four points for an A, three points for a B, two points for a C, one grade point for a D, and zero points for an F in each core academic course taken during high school (Garton et al., 2002). Core GPA is calculated by taking the total number of points received from core content courses, dividing by the total number of possible credits, then rounding to the nearest hundredth (Garton et al., 2002).

Composite ACT score. According to ACT (2014), composite ACT scores are scaled from 1 (lowest) to 36 (highest) regardless of the test version that is administered. A composite ACT score is calculated by summing the four sub-test (English, mathematics, reading, and science) scores, dividing by four, and then rounding to the nearest whole number (ACT, 2014).

English ACT score. According to ACT (2014), English ACT scores are scaled from 1 (lowest) to 36 (highest) regardless of the test version that is administered. The English section of the ACT measures student ability within three categories: production of writing, knowledge of language, and conventions of Standard English (ACT, 2014).

Mathematics ACT score. According to ACT (2014), mathematics ACT scores are scaled from 1 (lowest) to 36 (highest) regardless of the test version that is administered. The mathematics section of the ACT measures student ability within seven categories: number and quantity, algebra, functions, geometry, statistics and probability, integration of ideas, and mathematical modeling (ACT, 2014).

Reading ACT score. According to ACT (2014), reading ACT scores are scaled from 1 (lowest) to 36 (highest) regardless of the test version that is administered. The reading section of the ACT measures student ability within three categories: key ideas and details, craft and structure, and integration of knowledge and ideas (ACT, 2014).

Rural. Geverdt (2015) defined rural areas as census-defined territories that are more than five miles but less than or equal to 25 miles from an urbanized area.

Science ACT score. According to ACT (2014), science ACT scores are scaled from 1 (lowest) to 36 (highest) regardless of the test version that is administered. The science section of the ACT measures student ability within three categories: interpretation of data, scientific investigation, and evaluation of models, inferences, and experimental results (ACT, 2014).

Suburban. Geverdt (2015) defined suburban areas as geographical locations outside of a principal city and with a population less than 100,000.

Urban. Geverdt (2015) defined urban areas as geographical locations inside a principal city with a population over 250,000.

Organization of the Study

This study is comprised of five chapters. The first chapter included an introduction, the background of the study, statement of the problem, purpose of the study, significance of the study, delimitations, assumptions, research questions, and definitions of key terms. Chapter 2 is a comprehensive literature review which includes a history of grading practices in the United States, the rise of standards-based grading, reliability of grades, the configuration of grades, the ability of grades to predict educational outcomes, history of the ACT, studies regarding the relationship between standardized test scores and GPA, and a summary. A description of the methodology used in this study, which includes the research design, selection of participants, measurement, data collection procedures, data analysis procedures, hypothesis testing, and limitations is included in Chapter 3. The fourth chapter is a presentation of the findings of the study, including descriptive statistics, the results of the hypothesis testing, and a summary. Finally, Chapter 5 contains a summary of the entire study including an overview of the problem, purpose statement and research questions, review of the methodology, and major findings, findings related to the literature, and a conclusion with implications for action, recommendations for future research, and concluding remarks.

Chapter 2

Review of the Literature

This study was conducted to examine the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs calculated from all English, mathematics, history, and science courses in grades 9 through 12 in high schools in three south-central Kansas school districts. An additional purpose of the study was to explore how gender, race, socioeconomic status, and high school community type affect the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs for students in grades 9 through 12 in three south-central Kansas districts. Chapter 2 contains an in-depth review of the literature, which includes the history of grading practices in the United States, the rise of standards-based grading, discourse on the reliability of grades, a review of the configuration of grades, the ability of grades to predict educational outcomes, the history of the ACT, an overview of other studies about the relationship between standardized test scores and GPAs, and a chapter summary.

History of Grading Practices in the United States

Schools in colonial America were originally established as social agencies to promote morality, character, and virtue. Literacy was driven by the need to read the Bible, and unlike the secular public education of today, the school was subservient to the church. Despite the separation of church and state, modern grading practices still reward what is deemed to be virtuous behavior and punishes students for not following classroom procedure (Vatterott, 2015). Unfortunately, this use of grades created a culture where compliance and effort are of higher value than pure learning.

The history of grading practices in higher education has set the stage for grading practices in K-12. In the early days of group education, student progress reports were orally presented to parents without a comparison system to other students (Guskey & Bailey, 2001). Then, in the mid-16th century, Cambridge University pioneered the notion of ranking students in performance categories as a matter of efficiency (Winter, 1993). In the 16th century, Cambridge University and Yale University developed a three-tier grading system with 25 percent of the grades at the top, 50 percent in the middle, and 25 percent at the bottom (Winter, 1993). This practice later evolved into the standard five-point grading scale used in grade point averages today when schools turned to written descriptions of student skills in arithmetic, reading, and penmanship in place of the oral reports (Curreton, 1971; Durm, 1993; Guskey & Bailey, 2001; Rugg, 1918; Schneider & Hutt, 2014). Almost a century later, Harvard University began using percentages to justify student groupings and determined that students in the lowest group failed the course (Durm, 1993). Mount Holyoke College, in 1897, first crafted a system that combined the percentage groupings from Harvard with descriptive adjectives accompanied by letters of the alphabet (Durm, 1993). While grades were a universal aspect of the education system at the turn of the 20th century, they were idiosyncratically assigned (Schneider & Hutt, 2014).

Compulsory attendance laws established in the early 20th century drastically altered the backdrop of K-12 education in the United States. In a span of merely 40 years (1870-1910) student attendance at public elementary and secondary schools grew by over 11,000,000 students (Kirschenbaum et al., 1971). As public education entities grappled with their exponentially increasing enrollment, they turned to the systems created by

Yale, Harvard, and Mount Holyoke as a mechanism for ranking and classifying the student population.

Research from Starch and Elliott (1912, 1913), Crooks (1933), Hill (1935), Smith and Dobbin (1960), and Kirschenbaum et al. (1971) unveiled significant discrepancies between English and geometry scores dependent on the grader, thus fueling the debate about whether grading should be norm-based or criterion-referenced. To resolve inconsistencies, educators turned to the new field of statistics for answers on how to measure and rank human characteristics (Pearson, 1930; Galton & Galton, 1998). Inspired by the developing field of statistics and new theories put forth by Galton in the field of psychometrics, educators began investigating how these ideas could be applied to grading (Brookhart et al., 2016; Galton & Galton, 1998). When IQ scores from a random group of children were shown to fit a standard normal distribution, the bell curve replaced the 100-point scale as the premier grading feature of the early 20th century (Curreton, 1971).

Meyer (1908) argued for five grade categories: excellent (top 3% of students), superior (next 22% of students), medium (middle 50% of students), inferior (next 22% of students), and failure (bottom 3% of students). A decade after Meyer's (1908) proposal, 56.25% of secondary educations systems were using Meyer's five-point scale, typically associated with letters A through F (Nicholson, 1917). Rugg (1918) argued

> Teachers' marks on observations of students' performance should reflect the ability-to-do and form a normal distribution. That is, the normal distribution should form a basis for checking the quality of grade that teachers assign. This approach reduces grading to determining the number of grading divisions and the

> number of students who should fall into each category. Thus, there is a shift from a decentralized and fundamentally haphazard approach to assigning grades to one that is based on scientific principle. (p. 701)

According to Meyer (1908) and Rugg (1918), grades were merely a method for ranking students, not for making decisions about student intelligence or achievement. Despite Meyer (1908) and Rugg (1918) both stating that grades should merely represent a student's ranking in comparison to other students in the same class, teachers, parents, and students continue to use grades to make decisions about student achievement levels.

Fitting grades to a bell curve reflects an inherent misunderstanding of the concept of normal distribution (Kelly, 2008). Kelly (2008) even went so far as to claim that it is unreasonable to assume grades will be distributed in the same way as a large, untreated group of people. Despite research to the contrary, K-12 educators began to develop a consensus that a 0-100 student assessment system based on the normal curve would lead to precision in assigned grades and provide evidence of rigorous coursework (Guskey, 2000, 2006; Kulick & Wright, 2008).

In the 1920s, teachers began to adopt broad category, percentage-based grading systems in response to studies on the reliability of grades (Kelly, 1914; Rugg, 1918). While most elementary schools maintained the written descriptions developed in the early 20th century, secondary schools began to prefer percentage classifications as a time-saving and cost-effective measure to combat the demands of the diverse student population and subject specific instruction (Farr, 2000; Guskey & Bailey, 2001). Despite the widespread adoption of the percentage-based A through F grading scale, variation persisted (Hill, 1935). The standardized classification system did not prevent discrepancies in the frequency of grade reports, grading methodology, and definition of mastery (Hill, 1935). By 1940, over 80% of school districts in the United States had adopted the A through F grading scale based on percentage classifications (Brookhart et al., 2016). “One could argue that this move to percentage grades eliminated the specific communication of what students knew and could do” (Brookhart et al., 2016, p. 5).

In 1963, both Carroll and Glaser brought forth the notion that student grades should not be based on the performance of their peers, but rather criterion-referencing testing based on subject-area domains. According to Glaser (1963) mastery of material taught should be the impetus for grading systems. Working from Carroll’s (1963) and Glaser’s (1963) models, Bloom (1971) crafted the master learning theory: the grade received in any course should be a combination of the quality of teaching, student perseverance, and time required to reach mastery (Guskey, 1985). The Civil Rights movement and Vietnam War shifted the educational landscape once again (Rojstaczer & Healy, 2012).

Compliance grading began in earnest in the late 1960’s as educators used grades as a method to keep the swelling student ranks in line and prevent young men from having to enter the military draft (Goodwin, 2011). The combination of the Civil Rights movement and the Vietnam War led to a national reexamination of social assessment, learning, and equity (Kohn, 1999). Behaviorism was the leading classroom management philosophy and its ideas began to permeate grading practices as educators used good grades as a mechanism for rewarding positive behavior (Kohn, 1999). Then, as the Vietnam War was met with increasingly more resistance, teachers faced additional pressure to not assign failing grades, thus preventing students from being subjected to the draft (Kohn, 1999; Rojstaczer & Healy, 2012). The convergence of major societal events in the 1960's led to dramatic swings in grading perspectives. As grades shifted from mere achievement to compliance rewards, U.S. educators unintentionally began the monumental shift from learning-based grades to the effort-based grades so prevalent today (Rojstaczer & Healy, 2012).

According to Rojstaczer and Healy (2012), the proportion of A and B grades assigned rose dramatically (over 15%) between 1965 and 1975. Concurrently, the proportion of D and F grades assigned dropped nearly 20% (Rojstaczer & Healy, 2012). Once the draft ended, grade proportions did not return to the pre-Vietnam War levels (Rojstaczer & Healy, 2012). Between 1991 and 2003, high school grade point averages (GPA) for the same ACT score increased 12.5% (ACT, 2005) giving rise to claims about the invalidity of grades and cries for more standardized testing. In the 25 years after the Vietnam War, a new view of students as consumers emerged resulting in what Love and Kotchen (2010) deemed the grade-leniency theory. Grade-leniency theory contends that teachers inflate grades through behavior-based assessments and the desire for student approval (Love & Kotchen, 2010).

The percentage-based A through F grading scale has remained the preferential student progress reporting system for almost 80 years in the United States. However, the 21st century has seen the rise of standards-based grading (SBG) in response to the claims of grade invalidity (Abbott, 2008; Edwards, 2009; Grindberg, 2014; McCann, 2003; O'Halloran & Gordon, 2014; Olsen, 1995; Walsh, 2010). SBG is a rapidly evolving student progress reporting system in which effort, work ethic, and behavior are reported separately from achievement scores, which are based solely on mastery of standards

(Creason, 2013; Grindberg, 2014; Guskey & Bailey, 2010; Marzano, 1996; Vatterott, 2015). Although most secondary schools have maintained the norm-referenced grades needed to accommodate the college admissions process, educational systems across the United States are transitioning to SBG (Brookhart et al., 2016).

The Rise of Standards-Based Grading

Unlike the traditional A through F grading system, SBG describes student performance as below basic, basic, proficient, or advanced (Brookhart, 2011; Melograno, 2007). SBG emphasizes relating student progress to grade-level, subject specific standards and reporting behavior and student work habits separately from a course grade (Guskey, 2009; Guskey & Bailey, 2010; Marzano & Heflebower, 2011; McMillan, 2009; Mohnsen, 2013; O'Connor, 2009; Scriffiny, 2008; Shippy, Washer, & Perrin, 2013; Wiggins, 1994). Initial assertions claim SBG will create educational reform and instructional improvement because of the high-quality information available to students, parents, and teachers regarding specific skills (Guskey & Bailey, 2001). However, other researchers have claimed that SBG is subject to the same reporting errors and misinterpretations as the traditional A through F grading scale (Cizek, 2000).

Across the United States, school districts and teachers have experienced a myriad of issues when implementing SBG (Clarridge & Whitaker, 1994). The key to successful implementation of SBG is understanding by the teachers, students, and parents and keeping behaviorism-based scores out of academic grades (Guskey & Bailey, 2001). While a majority of teachers report that they back SBG reform and removing behaviorism scores, they simultaneously self-report continuing to mix work ethic, behavior, and tardy grades with academic scores in SBG systems (Cox, 2011; Hay &

McDonald, 2008; McMunn, Schenck, & McColskey, 2003). According to Cox (2011), late work policies, use of common assessments, and retake policies vary greatly from teacher to teacher, thus hindering school-based SBG efforts. McMunn et al. (2003) demonstrated that changes in grading policies do not necessarily result in changes to grading practice. Hay and McDonald (2008) and Cox (2011) reaffirmed the initial findings from McMunn et al. (2003). Simon et al. (2010) identified another obstacle in the implementation of SBG, tension between teacher practice and SBG ideals. In 2011, Tierney, Simon, and Charland expanded on the Simon et al. (2010) study and found that even teachers with extensive training in SBG were failing to apply fundamental SBG concepts to their actual grading practices.

Despite extensive recent literature on the merits and implementation of SBG, there are few empirical studies (Brookhart et al., 2016). Swan, Guskey, and Jung (2014) surveyed 115 parents and 383 teachers from a district that concurrently generated a traditional and an SBG report card. Both parents and teachers preferred the SBG report card due to the quality of information provided (Swan et al., 2014). Teachers from the Swan et al. study (2014) also commented that SBG report cards took substantially more time to generate. Swan et al.'s (2014) results mirror the smaller version of the same study completed by Guskey, Swan, and Jung in 2010. However, the Guskey et al. (2010) and Swan et al. (2014) research contradicts a Guskey (2004) report that found nearly 80% of parents attempted to interpret the below basic, basic, proficient, and advanced labels from SBG in terms of traditional letter grades.

Proponents of SBG claim that SBG is a more accurate record of student academic capacity. Therefore, one might expect a strong correlation between SBG GPAs and standardized assessment scores. However, all four research studies (Howley, Kusimo, & Parrott, 1999; Ross & Kostuch, 2011; Welsh & D'Agostino, 2009; Welsh, D'Agostino, & Kaniskan, 2013) indicated only a moderate correlation between SBG GPAs and standardized assessment scores. Howley et al. (1999) determined that half the variance in GPA can be determined by non-cognitive factors like attendance, behavior, and participation, despite the premise that SBG should not include non-cognitive factors. As expected, the strength of the correlation between SBG GPAs and standardized assessment scores varied by school (Howley et al., 1999). Ross and Kostuch (2011) conducted an extensive study of the relationship between SBG GPAs and standardized test scores with a random sample of 15,942 high school students in Ontario, Canada. Once again, the correlation was moderate (r = .61), but the relationship was not affected by student gender or grade level (Ross & Kostuch, 2011). Ross and Kostuch (2011) determined the correlation was stronger for mathematics (r = .63) than for reading (r = .59) or writing (r =.58) and grades tended to be higher than expected in all subject areas.

Welsh and D'Agostino (2009) established an Appraisal Style Scale to fully assess the use of high-quality SBG practices and their relationship to standardized assessments. Welsh and D'Agostino (2009) found a moderate correlation (r = .49) in the convergence rate between Appraisal Style Scale score on a 10-point scale and standardized assessment score. In 2013, Welsh et al. replicated the 2009 study and determined that SBG GPAs and standardized assessment scores were stronger in mathematics than in reading, despite students receiving their lowest grades in mathematics. Overall, researchers have found little evidence to support the notion that SBG GPAs are more related to standardized assessment scores than teachers who use the traditional A through F grading scale

(Howley et al., 1999; Ross & Kostuch, 2011; Welsh & D'Agostino, 2009; Welsh et al., 2013). Regardless of type of GPA (traditional or SBG), researchers have found at most a moderate correlation between students' standardized assessment score and GPAs (Bolek, 2011; Ginstead, 2013; Howley et al., 1999; Jones, 2008; Mahlum, 2015; Noble, Davenport & Sawyer, 2001; Noble & Schnelker, 2007; Roberts & Noble, 2004; Ross & Kostuch, 2011; Roth, Crans, Carter, Ariet, & Resnick, 2000; Schiel, Pommerich, & Noble, 1996; Welsh & D'Agostino, 2009; Welsh et al., 2013).

Reliability of Grades

Research on the reliability of grades can be traced back to the late 19th century. Edgeworth (1888) designated three sources of error in teacher-assigned grades: chance, personal differences among graders and exams, and understanding of proficiency. Furthermore, Edgeworth (1888) identified potential consequences of grading variability and used his research to advocate for improving the reliability of grades. Edgeworth's (1888) research propelled a litany of other studies on the variation of grades in the early 20th century (Ashbaugh, 1924; Bolton, 1927; Eells, 1930; Healy, 1935; Hulten, 1925; Jacoby, 1910; Kelly, 1914; Lauterbach, 1928; Rugg, 1918; Shriner, 1930; Silberstein, 1922; Sims, 1933; Starch, 1913, 1915; Starch & Elliot, 1912, 1913). Of the 16 listed studies, 13 researchers determined that there was significant variation in the grades teachers assign to the same student work (Ashbaugh, 1924; Eells, 1930; Healy, 1935; Hulten, 1925; Kelly, 1914; Lauterbach, 1928; Rugg, 1918; Silberstein, 1922; Sims, 1933; Starch, 1913, 1915; Starch & Elliot, 1912, 1913).

As previewed by Edgeworth (1888), differences or a lack in grading criteria was consistently found to be a prominent source of grading variability (Ashbaugh, 1924;

Eells, 1930, Healy, 1935; Silberstein, 1922; Starch, 1913). Teacher inconsistency over time was also a major source of variability (Eells, 1930; Hulten, 1925; Silberstein, 1922). Even when graded by the same teacher a second time, over 20% of the grades were changed from pass to fail (Hulten, 1925; Silberstein, 1922). Eells (1930) determined a weak intra-teacher reliability correlation (r = .25 to .51). Starch (1913, 1915) and Starch and Elliott (1913) conducted the most extensive early research on the reliability of grades, with 414 teachers grading 38 compositions or exams. Starch and Elliott (1912, 1913) found teacher assigned grades to vary by as many as 40 points on a 100-point scale. Variability was highest in grading history exams, with a standard change of 7.7% (Starch & Elliott, 1913). The variability of history grades is followed closely by mathematics (standard change of 7.5%) and English (standard change of 4.8%) (Starch & Elliott, 1913). Starch (1913) isolated four sources of variation and reported the probable error on a 100-point scale:

> 1) Differences among the standards of different schools (probable error almost 0),
>
> 2) Differences among the standards of different teachers (probable error = 1.0),
>
> 3) Differences in the relative values placed by different teachers upon various elements in a paper, including content and form (probable error = 2.1), and
>
> 4) Differences due to the pure inability to distinguish between closely allied degrees of merit (probable error = 2.2). (p. 632)

As a result of his research, Starch (1913) recommended the use of the nine-point A through F plus-minus scale as opposed to the five-point regular A through F scale to eliminate variability. Later, Starch (1915) tested his recommendation and found grade variability was reduced by 4.2% when using the nine-point scale instead of the five-point scale. Sims (1933) echoed Starch's (1915) results when he determined two kinds of variability in teacher assigned grades: "(a) differences in students' work quality and (b) differences in the standards of grading found among school systems and among teachers within a system" (p. 637).

Bolton (1927), Jacoby (1910), and Shriner (1930), conducted three other early 20th century studies and contended that while there was variability in teacher grading, the variability was not as great as Ashbaugh (1924), Hulten (1925), Kelly (1914), Rugg (1918), Silberstein (1922), Starch (1913, 1915), and Starch & Elliot (1912, 1913) suggested. Jacoby (1910) conducted a small descriptive statistics study with six different teachers grading 11 essays and determined that his low variability was an indication of the high quality of student work. Bolton (1927) and Shriner (1930) reaffirmed Jacoby's (1910) conclusion when they both found greater grade variability in low quality student work than high quality papers. Shriner (1930) also determined that teacher grades were reliable since the median correlations of each teacher's grade with the average grade of the paper were r =.946 in Algebra and r = -.917 in English. Shriner's (1930), Bolton's (1927), and Jacoby's (1910) results were later refuted by Eells (1930) who found greater grading consistency in poorer papers. The quality of student work affected the variability in teacher assigned grades, but the results from various studies were inconsistent.

Research on grading reliability in late 20th and early 21st century has focused on the unintended adverse consequences of variability and the multidimensional factors of academic knowledge, persistence, and substantive engagement that impact grades (Messick, 1989). The earliest study involving non-cognitive variables was from Miner (1967), who analyzed longitudinal data for a sample of 671 high school students. She identified that in addition to objective achievement on standardized tests, student grades were affected by attendance, punctuality, effort, and behavior (Miner, 1967). In 1990, Farkas, Grobe, Sheehan, and Shaun reiterated Miner's (1967) result by showing that student work habits were the strongest non-cognitive predictors of grades in their sample of 486 eighth-graders, noting "teacher judgments of student non-cognitive characteristics are powerful determinants of course grades, even when student cognitive performance is controlled" (p. 140). Similarly, Willingham, Pollack, and Lewis (2002) identified a strong positive relationship (r = .88) between student grades and student engagement, persistence, motivation, and work completion. Willingham et al. (2002) found only a moderate relationship between grades and ACT scores (r = .51) for 8,454 students from 581 different high schools, suggesting grades are a more useful assessment for non-cognitive than cognitive student factors.

Kelly (2008) used a hierarchical linear modeling study of 1,653 middle school students to argue that while grades contain non-cognitive factors, doing so leads to increased engagement and effort, and thus increased learning. Student grades were significantly, positively impacted by non-cognitive factors, suggesting "most teachers successfully use grades to reward achievement-oriented behavior and promote a widespread growth in achievement" (Kelly, 2008, p. 45). Kelly (2008) also argued that despite his results, high-stakes standardized testing is not inherently more objective than teacher-written tests. Brimi (2011) countered Kelly's (2008) conclusion when 73 English teachers grading one essay led to a score range of 46 points on a 100-point scale and encompassed all five grades, A through F. Such a substantial range of results led Brimi (2011) to advocate for the use of national standardized assessments.

Recent studies have expanded Kelly's (2008) and Brimi's (2011) research with enlarged sample sizes and sophisticated structural equation modeling and multi-dimensional scaling studies. Bowers (2009, 2011) examined the relationship between grades and ACT scores in core subjects (English, science, mathematics, and history) and non-core subjects (business, art, foreign language, and physical education). Three factors in the reliability of grades emerged: "a cognitive factor that describes the relationship between tests and core subject grades, an engagement factor between core subject grades and non-core subject grades, and a factor that describes the difference between grades in art and physical education" (Bowers, 2011, p. 154). Overall, Bowers (2011) argued that a majority of the variance in grades not related to test scores stems from teachers' assessment of students' ability to meet the social processes within a school's structure.

Casillas et al. (2012) examined the relationship between GPAs and standardized assessment scores of 4,660 students. Academic commitment and student success factor (degree of student conscientiousness and effort) had the strongest correlation to GPA (r = .84) (Casillas et al., 2012). On a 100-point scale, variance in GPA was attributed to each of the following factors: combination of prior grades (30%), standardized assessment scores (24%), psychosocial factors (23%), student behavior indicators (10%), demographic data (9%), and school circumstances (3%) (Casillas et al., 2012).

The findings of Kelly (2008), Bowers (2009, 2011), and Casillas et al. (2012) support the notion that tested achievement only represents one dimension of schooling and is distinct from the skills measured by non-cognitive student factors (Resnick, 1987). "A validity argument that expects grades and standardized tests to correlate highly therefore may not be sound because the construct of school achievement is not fully defined by standardized test scores" (Brookhart et al., 2016, p. 16). Thus, researchers began to shift the focus of their inquiry from the reliability of grades to how grades are configured.

Configuration of Grades

The educational experience of all American school children is centered on grades, therefore grading research is an integral aspect of our understanding of school success (Brookhart et al., 2016). In fact, some of the earliest educational research focused on grading and what grades mean (Brookhart, 1993). In over 100 years of educational research, there have been numerous claims about the inability of grades to measure academic achievement (Allen, 2005; Banker, 1927; Carter, 1952; Evans, 1976; Hargis, 1990; Kirschenbaum et al., 1971; Quann, 1983; Simon & Bellanca, 1976). Other researchers have noted that grades are a powerful performance indicator for schools, teachers, students, and parents (Bisesi, Farr, Greene, & Haydel, 2000; Folzer-Napier, 1976; Linn, 1982). Recently, researchers have attempted to capture the various components of a grade to inform decision-making and predict the completion of a secondary and post-secondary education (Atkinson & Geiser, 2009; Bowers, 2009; Bowers, Sprott, & Taff, 2013). This research leads to questions about whether grades are meant to rank students based on academic knowledge, predict future outcomes, motivate student performance, or a combination of all three. Over 100 years of educational research has shown that grades typically represent a mixture of what a particular teacher values: cognitive knowledge, engagement, persistence, normed social behavior, etc. (Allensworth, 2013; Allensworth & Easton, 2007; Allensworth, Gwynne, Moore, & de la Torre, 2014; Aronson, 2008; Atkinson & Geiser, 2009; Bowers, 2009, 2011, 2014; Cizek, Fitzgerald, & Rachor, 1995; Cross & Frary, 1999; Duncan & Noonan, 2007; Farkas et al., 1990; Guskey, 2002, 2009; Imperial, 2011; Kelly, 2008; Klapp Lekholm & Cliffordson, 2008, 2009; Liu, 2008; McMillan, 2001; McMillan & Lawson, 2001; McMillan, Myran, & Workman, 2002; McMillan & Nash, 2000; Miner, 1967; Randall & Engelhard, 2009, 2010; Russell & Austin, 2010; Sun & Cheng, 2013; Svennberg, Meckbach, & Redelius, 2014; Troug & Friedman, 1996; Willingham et al., 2002; Yesbeck, 2011).

More than 60% of teachers endorse academic enhancers like work ethic, participation, and effort as important factors in achievement and motivation, and thus deem them as relevant to a student's course grade (Aronson, 2008; Cizek et al., 1995; Cross & Frary, 1999; Duncan & Noonan, 2007; Guskey, 2002, 2009; Imperial, 2011; Kelly, 2008; Liu, 2008; McMillan, 2001; McMillan & Lawson, 2001; McMillan, Myran, & Workman, 2002; McMillan & Nash, 2000; Randall & Engelhard, 2009, 2010; Russell & Austin, 2010; Sun & Cheng, 2013; Svennberg et al., 2014; Troug & Friedman, 1996; Yesbeck, 2011). Therefore, grades cannot be thought of as one-dimensional measures purely based on academic skills, as presumed in the past (Brookhart, 2009, 2011; Carter, 1952; Guskey, 2000; Guskey & Bailey, 2010; Marzano & Hefflebower, 2011; McCandless, Roberts, & Starnes, 1972; Moore, 1939; O'Connor, 2009; Ross & Hooks, 1930; Scriffiny, 2008). While researchers may hope that grades are unsullied measures of academic progress, strong evidence indicates that they are actually multidimensional measures that simultaneously reflect academic achievement, student effort, and behavioral expectations (Brookhart, 2009, 2011; Guskey, 2000; Guskey & Bailey, 2010; Marzano & Hefflebower, 2011; O'Connor, 2009; Scriffiny, 2008). The elements of education a teacher values typically plays a major role in how the teacher assigns grades.

In addition to tested academic skills and academic enhancers, most teachers place high value on being fair (Bonner, 2016; Bonner & Chen, 2009; Brookhart, 1994; Grimes, 2010; Hay & MacDonald, 2008; Sun & Cheng, 2013; Svennberg et al., 2014; Tierney et al., 2011). The high value placed on fairness is characteristically displayed by the notion that students who work hard and try should not be subjected to failing grades (Bonner, 2016; Bonner & Chen, 2009; Brookhart, 1994; Grimes, 2010; Hay & MacDonald, 2008; Sun & Cheng, 2013; Svennberg et al., 2014; Tierney et al., 2011). Concepts such as academic enhancers, fairness, and effort, skew the grading paradigm away from purely academic measures and result in high variability in teacher assigned grades (Ashbaugh, 1924; Bolton, 1927; Brimi, 2011; Eells, 1930; Healy, 1935; Hulten, 1925; Jacoby, 1910; Lauterbach, 1928; Shriner, 1930; Silberstein, 1922; Sims, 1933, Starch, 1915; Starch & Elliott, 1913).

Research on teacher assigned grades has attributed the variability to six major causes: assigned task, work quality, grading scale, criteria, instructor leniency, and teacher error (Ashbaugh, 1924; Bolton, 1927; Brimi, 2011; Eells, 1930; Healy, 1935; Hulten, 1925; Jacoby, 1910; Lauterbach, 1928; Shriner, 1930; Silberstein, 1922; Sims, 1933, Starch, 1915; Starch & Elliott, 1913). According to Starch and Elliot (2013), criteria and teacher error are the two most significant sources of grade variability. Beatty, Walmsley, Sackett, Kuncel, and Koch (2015) determined the reliability of college GPAs to be a strong .93, but secondary GPAs only carried a reliability score of .74 when grades were aggregated from individual work. However, Myford (2012) showed that teacher assigned grades were more accurate and reliable when the teachers received training on established criteria and graded in groups instead of individually. Collective grading is a rare practice but appears to be a promising method to enhance the reliability of grades, in conjunction with clearer criteria and involving students in rubric development (Myford, 2012).

The criteria for grades construct the function of grades and what grades measure. In the 19th and 20th centuries, researchers postulated that when teachers included non-cognitive factors that did not indicate cognitive skill, grades failed to give administrators, parents, students, and community members an accurate indication of school, teacher, and student success (Brookhart, 1994). Despite the push for grades based only on cognitive ability, there is extensive evidence that teachers also included student improvement, effort, behavior, and other academic enhancers when assigning grades (Atkinson & Geiser, 2009; Cox, 2011; Hay & McDonald, 2008; McMunn et al., 2003). Current theories postulate that the non-cognitive dimension of grades is critical to students' social, emotional, and character development (Cunha & Heckman, 2008; Heckman & Rubinstein, 2001; Levin, 2013). While non-cognitive skills can aid students in their acquisition of cognitive concepts, cognitive concepts do not aid students in non-cognitive development (Cunha & Heckman, 2008). Since grades are multidimensional measures of student achievement that may explain why grades are more accurate predictors of educational outcomes than standardized test scores (Atkinson & Geiser, 2009; Barrington & Hendricks, 1989; Bowers, 2014; Cairns, Cairns, & Neckerman, 1989; Cliffordson, 2008; Ekstrom, Goertz, Pollack, & Rock, 1986; Ensminger & Slusarcick, 1992; Finn, 1989; Fitzsimmons, Cheever, Leonard, & Macunovich, 1969; Hargis, 1990; Lloyd, 1974, 1978; Morris, Ehren, & Lenz, 1991; Nichols & Berliner, 2007; Polikoff, Porter, & Smithson, 2011; Rumberger, 1987; Troob, 1985; Voss, Wendling, & Elliott, 1966).

Using Grades to Predict Educational Outcomes

While grades are alleged to be the reward for academic achievement and used to grant high school diplomas (Rumberger, 2011), they also permeate the school experience and inform educational outcomes (Brookhart, 1993; Pattison, Grodsky, & Muller, 2013). In a late 20th and early 21st century vein of educational research regarding grades, researchers studied the relationship between grades and educational outcomes (Alexander, Entwisle, & Kabbani, 2001; Allensworth & Easton, 2007; Allensworth et al., 2014; Balfanz, Herzog, & MacIver, 2007; Barrington & Hendricks, 1989; Bowers, 2010; Bowers & Sprott, 2012; Bowers et al., 2013; Cairns et al., 1989; Cliffordson, 2008; Ekstrom et al., 1986; Ensminger & Slusarcick, 1992; Fitzsimmons et al., 1969; Jimerson et al., 2000; Lloyd, 1978; Morris et al., 1991; Roderick & Camburn, 1999; Troob, 1985) and offered evidence of the validity of grades (Brookhart, 2015). Grades predict important educational consequences such as dropping out of school (Bowers, 2010; Bowers & Sprott, 2012; Bowers et al., 2013), matriculating to college (Atkinson & Geiser, 2009; Bowers, 2010; Thorsen & Cliffordson, 2012), and successful completion of a higher education (Atkinson & Geiser, 2009; Bowers, 2010; Sawyer, 2013; Thorsen & Cliffordson, 2012).

Early studies regarding grades and educational outcomes determined teacher assigned grades were the strongest predictor of failure to graduate from high school (Fitzsimmons et al., 1969; Lloyd, 1978; Voss et al., 1966). Research from the late 20th and early 21st centuries, demonstrated the cumulative effect of teacher assigned grades on a student’s decision to drop out (Alexander et al., 2001; Gleason & Dynarski, 2002; Jimerson, Egeland, Sroufe, & Carlson, 2000; Pallas, 2003; Roderick & Camburn, 1999).

Further research added student absence count and disciplinary action as strong predictors of student risk for failure to graduate from high school (Barrington & Hendricks, 1989; Cairns et al., 1989; Ekstrom et al., 1986; Ensminger & Slusarcick, 1992; Finn, 1989; Hargis, 1990; Morris et al., 1991; Rumberger, 1987; Troob, 1985).

Current research on the ability to predict educational outcomes from GPA has held a two-pronged focus: longitudinal data used to measure the influence of Ds and Fs (Allensworth & Easton, 2007; Balfanz et al., 2007; Bowers & Sprott, 2012; Bowers et al., 2013) and GPA in comparison to standardized test scores (Allensworth et al., 2014; Bowers, 2010; Cliffordson, 2008). Allensworth and Easton (2007) examined data regarding 24,894 Chicago ninth graders and determined that failing a core subject (English, mathematics, history, or science) was significantly correlated to dropping out of high school since it derailed a student from being on-track to graduate. Balfanz et al. (2007) replicated Allensworth and Easton's (2007) results with 12,972 sixth graders from Philadelphia. In 2012, Bowers and Sprott found that having a D or an F in a core course was second only to grade retention as the strongest predictor of dropping out. The next year, Bowers et al. (2013) determined that GPA was the most accurate dropout criterion.

Cliffordson (2008) conducted the most extensive study comparing GPA and standardized test scores as educational predictors. After reviewing data about 164,106 students, Cliffordson (2008) determined that secondary school GPA more accurately predicted achievement in higher education than students' norm-referenced score on the Swedish Scholastic Aptitude Test (SweSAT). Bowers (2010) replicated Cliffordson's (2008) study on a much smaller scale in the United States. Low grade clusters across core subjects (English, mathematics, history, and science) and being retained in a grade had a strong correlation with dropping out, while ACT scores failed to create a correlation. While Cliffordson (2008) studied the transition from secondary to higher education, Allensworth et al. (2014) found that similar trends affected the transition from middle school to high school. In 19,963 Chicago eighth-graders, grades and attendance were stronger predictors of high school academic performance than the standardized test scores (Allensworth et al., 2014). While the intention of grades is to assign a rating to a student's cognitive skill, grades have consistently shown to be accurate educational predictors (Bowers, 2014).

History of the ACT

The Standardized Achievement Test (SAT), was created in 1923 to measure academic achievement and preparation for college (ACT, 2016). In 1959, Lindquist founded the ACT with the goal of focusing on cognitive skills taught in secondary school instead of cognitive reasoning like the SAT (ACT, 2016). Over time, the landscape of the United States' student population evolved, and it became popular for students on the coasts to take the SAT and students in the middle of the country to take the ACT to meet college entrance requirements (ACT, 2016). However, the number of students registered for the ACT did not surpass the number of students registered for the SAT until 2012 (ACT, 2016). In 2013, ACT rearranged the reading and mathematics portion of the test to better align the test to the National Board's Common Core Standards (ACT, 2016). The reading section of the test removed three passage detail questions and replaced them with the integration of ideas questions (ACT, 2016). Additionally, the mathematics section of the test added two statistics and probability questions after removing two trigonometry questions (ACT, 2016). For the last six years, ACT has remained the leading college admissions test, leading to a strong push from school districts to raise student scores (ACT, 2016).

The prevalence of the ACT in college admissions and scholarship dollars has led to increased school district awareness and action regarding the test (ACT, 2015; McDaniel, 2015). As shown in Table 2, Kansas students have traditionally averaged one point higher than the national average on ACT subsection and composite scores (ACT, 2017b).

Table 2

State and National Average ACT Scores including Subsections 2013-2017

School Year	English		Mathematics		Reading		Science		Composite	
	KS	US	KS	US	KS	US	KS	US	KS	US
2013	21.2	20.2	21.7	20.9	22.3	21.1	21.7	20.7	21.8	20.9
2014	21.4	20.3	21.7	20.9	22.5	21.3	21.8	20.8	22.0	21.0
2015	21.3	20.4	21.6	20.8	22.4	21.4	21.8	20.9	21.9	21.0
2016	21.3	20.1	21.5	20.6	22.5	21.3	21.8	20.8	21.9	20.8
2017	21.1	20.3	21.3	20.7	22.3	21.4	21.7	21.0	21.7	21.0

Note. Adapted from *The Condition of College and Career Readiness 2017 – Profile Report – Kansas,* by ACT, 2017b, p. 7. Retrieved from http://www.act.org/content/dam/act/unsecured/ documents/cccr2017 /P_17_179999_S_S_N00_ACT-GCPR_Kansas.pdf

Despite ACT scores consistently higher than the national average, 27% of Kansas students are required to enroll in remedial post-secondary coursework (McDaniel, 2015). Failure to meet core course requirements explains the high percentage of Kansas students who need to take remedial courses despite the state's high ACT average. For data purposes ACT (2017b) has defined meeting core course requirements as completing four or more years of English and three or more years of mathematics, social studies, and natural science. Table 3 delineates Kansas ACT data for students who have met or exceeded core course requirements and students who have not.

Table 3

Kansas Average ACT Scores 2013-2017 for Core and Non-Core

School Year	English		Mathematics		Reading		Science		Composite	
	Met	Not Met	Met	Not Met	Met	Not Met	Met	Not Met	Met	Not Met
2013	22.0	18.0	22.4	18.9	22.9	19.6	22.4	19.3	22.6	19.1
2014	22.2	18.5	22.3	19.1	23.2	20.0	22.5	19.5	22.7	19.4
2015	22.1	18.2	22.2	19.0	23.1	19.8	22.4	19.4	22.6	19.2
2016	22.1	18.5	22.1	19.1	23.2	20.1	22.4	19.7	22.6	19.5
2017	21.9	18.6	21.9	19.1	23.1	20.1	22.3	19.6	22.5	19.5

Note. Adapted from *The Condition of College and Career Readiness 2017 – Profile Report – Kansas,* by ACT, 2017b, p. 8. Retrieved from http://www.act.org/content/dam/act/unsecured/ documents/cccr2017 /P_17_179999_S_S_N00_ACT-GCPR_Kansas.pdf

Students who have met the core course requirements as defined by ACT (2017b) typically score two to three points higher on every subsection of the ACT than students who have not met the core course requirements. School districts have begun to tackle the high need for their graduates to enroll in remedial courses by embedding ACT preparation and standard study skills across curricula (ACT, 2017a). More school districts subscribed to ACT Aspire, ACT Plan, PreACT, and ACT preparation resources in 2017 than in the previous 10 years combined (ACT, 2017a). As community pressure for positive ACT results mount, school districts have funneled money into preparing students for the test and raising student academic profiles (ACT, 2017a).

Researchers (Bolek, 2011; Ginstead, 2013; Jones, 2008; Mahlum, 2015; Noble, Davenport & Sawyer, 2001; Noble & Schnelker, 2007; Roberts & Noble, 2004; Roth et al., 2000; Schiel et al., 1996) have conducted numerous studies about the relationship between the ACT and coursework or GPA over the last two decades, including how gender, race, school classification, and socioeconomic status affect the relationship. The research has shown, with few exceptions, that a student's GPA is not related to the composite ACT score, while factors such as teacher, curricula, courses taken, and school classification are related (Bolek, 2011; Ginstead, 2013; Jones, 2008; Mahlum, 2015; Noble et al., 2001; Noble & Schnelker, 2007; Roberts & Noble, 2004; Roth et al., 2000; Schiel et al., 1996).

In 1996, Schiel et al. used statistically controlled ACT Plan scores to determine the relationship between mathematics and science ACT scores and type of mathematics and science courses taken. Schiel et al. (1996) used the ACT Plan test to control for prior student achievement since the ACT Plan is a pre-ACT test that predicts students' future success on the ACT. Junior and senior students who took advanced honors or advanced placement (AP), science and mathematics courses had significantly higher ACT scores than students who did not (Schiel et al., 1996). Student scores were not affected by gender, race, or socioeconomic status, and were not related to ACT Plan scores (Schiel et al., 1996). Roth et al. (2000) confirmed Schiel et al.'s results when their extensive review of Florida high school student records revealed that enrollment in advanced coursework, not necessarily passing the advanced classes, yields higher ACT scores and an easier post-secondary transition. "Exposure to the coursework helped prepare students to pass [the test], which nullifies the notion that students with high grades will automatically pass the exam" (Roth et al., 2000, p. 73). Roth et al.'s (2000) research supports the notion of enrolling students in advanced coursework for exposure, even if the student does not receive a high grade in the course. Roberts and Noble (2004) further established the importance of selected coursework, when their review of 175,000 ACT Plan scores from rural Iowa high school sophomores showed that students who were enrolled or planned to enroll in Geometry, Algebra 2, Trigonometry, or Chemistry courses consistently scored higher than students who did not plan to take such courses, regardless of academic history, gender, race, or socioeconomic status. Noble and Schnelker (2007) continued Roberts and Noble's (2004) work and studied student enrollment in English courses as well. The extensive study, with 403,381 students who took the ACT Plan as sophomores and the ACT as juniors and seniors, utilized 10,792 high schools that used sections from the ACT to develop portions of the curricula (Noble & Schnelker, 2007). Similar to the previous studies, Noble and Schnelker (2007) found that advanced coursework, not GPA was the most accurate indicator of future ACT score. Additionally, Noble and Schnelker (2007) noted that the high school classification, percent of qualified teaching staff, and school district access to ACT preparation resources all affected the correlation coefficient between GPA and ACT score.

Exploring the relationship between GPA and ACT score has become an increasingly popular dissertation topic in the 21st century. Studies conducted in Wisconsin (Jones, 2008), Colorado (Bolek, 2011), Iowa (Ginstead, 2013), and Wyoming (Mahlum, 2015) have consistently found a moderate correlation between GPA and ACT score, and a strong correlation between selected coursework and ACT score. Jones (2008) used data from one high school in Wisconsin to determine the relationship between student coursework in mathematics, English, science, history, GPA, gender, and ACT score. According to Jones (2008), neither student GPA nor gender have a strong correlation to ACT score. However, selected content area coursework (English, mathematics, history, and science) was a strong indicator of student success on the ACT (Jones, 2008). Bolek (2011) replicated Jones' (2008) study with a focus on core GPA and ACT score for Colorado high school students. The "relationship results for the comparison of classroom grades to the Colorado Student Assessment Program (CSAP) demonstrated a strong positive correlation in English, mathematics, and science" (Bolek, 2011, p. 3), but the data did not indicate a relationship between classroom grades and ACT. Ginstead's (2013) research focused on mathematics grades and ACT mathematics scores of Iowa seniors in the class of 2012. While Ginstead's (2013) results affirm Jones' (2008) conclusion about the relationship between advanced coursework and ACT score, he noted the relationship was weaker for female and minority students than male and Caucasian students. Mahlum (2015) took a unique approach to the question of the relationship between high school students' GPA and ACT score by collecting scholarship data from students at Wyoming College for 12 consecutive years. Despite the scholarship being based on ACT score and GPA, Mahlum (2015) observed that student GPA and ACT scores remained flat. Despite a lack of support from previous studies, people continue to assume a strong positive correlation between grades and ACT scores.

Relationship between Standardized Assessment Scores and GPAs

Numerous studies have been conducted to investigate the relationship between standardized test scores and grades. As stated by Willingham, Pollack, and Lewis (2002):

> In some educational contexts we use tests to keep grade scales honest…because we do not fully understand or trust grades to be an accurate indicator of

> [standardized test] outcomes. But we also reverse those sentiments and use grades to demonstrate the usefulness of [standardized] tests and justify their use …the tendency is to assume that a grade average and a test score are…mutual surrogates; that is, measuring the same thing. (p. 2)

Typically, the studies regarding GPA and standardized test scores are "an effort to understand the configuration of grades and marks that teachers assign to K-12 students" (Brookhart et al., 2016, p. 813). Despite an assumption that there is a strong positive correlation between standardized test scores and grades, the correlation has consistently been a moderate positive correlation around $r = .50$ (Allen, 2005; Bowers, 2011; Brennan, Kim, Wenz-Gross, & Siperstein, 2001; Carter, 1952; Duckworth, Quinn, & Tsukayama, 2012; Duckworth & Seligman, 2006; Linn, 1982; McCandless et al., 1972; Moore, 1939; Pattison et al., 2013; Ross & Hooks, 1930; Stanley & Baines, 2004; Unzicker, 1925; Woodruff & Ziomek, 2004).

Research on the relationship between grades and standardized assessment scores can be delineated into three key periods: pre-1950, 1950-2000, and 2000-present. Early 20th century researchers compared grades to intelligence test scores. Unzicker (1925) found a correlation coefficient of $r = .47$ when comparing students' average grades across English, mathematics, and history to scores on the Otis intelligence test. In 1930, Ross and Hooks completed a meta-analysis of 20 studies conducted between 1920 and 1929 on intelligence test scores and report card grades for elementary, junior high, and high school students. With correlation coefficients ranging from $r = .38$ to .44, Ross and Hooks (1930) concluded the meta-analysis revealed grade-based records are "the most usable or practical of all bases for prediction of future success" (p. 195). Additionally, Ross and Hooks (1930) noted the surprisingly low correlation indicates that grades are not an indication of intellectual capacity, but intellectual compliance. Moore (1939) conducted a correlation study between grades and Stanford Intelligence Test scores of 200 middle school students with $r = .61$. As early as 1947, Swineford noted "the data [in the study] clearly show that marks assigned by teachers are reliable measures of something but there is apparently a lack of agreement on just what that something should be" (p. 47). Early 20th century researchers, particularly Swineford (1947), recognized that the moderate correlation between grades and intelligence test scores indicated a discrepancy between what the variables were measuring.

From 1950-2000, researchers moved from correlating grades and intelligence test scores to correlating grades with standardized assessment measures with a focus on criterion validity (Brookhart et al., 2016). Carter (1952) correlated the grades and standardized algebra achievement scores of 235 high school students, $r = .52$. McCandless et al. (1972) expanded Carter's (1952) study to 433 students and factored in socio-economic status, race, and gender as confounding variables ($r = .31$). Despite the 50-year time span, relatively little comprehensive research was conducted to study the relationship between teacher-reported grades and standardized achievement results. Both major studies conducted between 1950 and 2000 (Carter, 1952; McCandless et al., 1972) represent the 21st century trend of an increased number of researchers attempting to determine how teacher-assigned grades correlate to known standardized assessments scores (Brookhart et al., 2016).

As the late 20th century transitioned to the early 21st century, researchers duplicated earlier studies with larger samples, current standardized tests, and modern sampling methods (Brennan et al., 2001; Duckworth et al., 2012; Duckworth & Seligman, 2006; Pattison et al., 2013; Woodruff & Ziomek, 2004; Zhang & Sanchez, 2013). Brennan et al. (2001) compared the Massachusetts standardized state reading test scores to mathematics, English, and science grades of 736 eighth-grade students and found correlations ranging from r =.54 to .59. Similar to Brennan et al. (2001), Duckworth and Seligman's (2006) correlation study of 140 eighth-grade students' GPA and TerraNova California Achievement Test found a correlation of r =.66, the highest correlation of any study comparing grades and standardized achievement test scores.

Woodruff and Ziomek (2004) undertook the most comprehensive study of the relationship of grades and standardized assessment scores. Woodruff and Ziomek (2004) compared composite ACT scores and GPA for every high school student who took the ACT between 1991 and 2003. The analysis, consisting of a sample size of nearly 700,000 unique students, found moderate positive correlations for self-reported average GPA and composite ACT scores (r = .56 to .58), self-reported mathematics grades and ACT mathematics scores (r = .54 to .57), and self-reported English grades and ACT English scores (r = .45 to .50) (Woodruff & Ziomek, 2004). "The results of the study indicated that students who took the extra year of mathematics, science, and social studies did better on the ACT than students who completed the minimum or less than the core class requirements" (Woodruff & Ziomek, 2004, p. 19). Zhang and Sanchez (2013) followed Woodruff and Ziomek's (2004) study comparing high school GPAs and ACT scores with divergent results. "From 1991 to 2003 there was an increase in [core] high school GPA for the selected ACT scores. After 2003, there was comparatively little change in the [core] high school GPA scores" (Zhang & Sanchez, 2013, p. 10).

As recently as 2012, reports released by the United States Department of Education (USDE) conveyed a paradox in student GPAs in science, mathematics, and reading courses and corresponding scores on federal standardized tests (USDE, 2012a, 2012b, 2012c). While student enrollment and GPAs in upper level courses across the United States are increasing, the U.S. has consistently scored lower than other similarly developed countries on international standardized tests, and U.S. scores have dropped every year since 1999 (Schmidt, 2007). The divergence between GPAs and standardized test scores is widening for students with a motivation or cognitive skill discrepancy according to the Office of Institutional Research and Policy Studies (OIRP) (OIRP, 2011). Duckworth et al. (2012) found results divergent from the USDE reports. In a structural equation modeling study of 1,364 ninth grade and 510 eighth-grade students, the correlation (r = .62 to .65) between reading and mathematics scores from the ACT Plan, a preliminary ACT assessment, and GPA was stronger than prior studies.

Pattison et al. (2013) conducted the second largest correlation study to compare high school GPA to reading and mathematics standardized test results. The researchers compiled a sample size of over 40,000 students spanning 1972 to 2002 by obtaining GPA and ACT content scores in reading and mathematics from the National Longitudinal Study of the High School Class of 1972, High School and Beyond sophomore cohort, National Educational Longitudinal Study of 1988, and Educational Longitudinal Study of 2002 (Pattison et al., 2013). The correlations ranged from r = .46 to .54 for reading and r = .52 to .64 for mathematics, consistent with the findings in past research.

Despite variability of years, subjects, systemic educational shifts, teacher accountability, and emphasis of standardized testing, “correlations have remained moderate but remarkably consistent in studies based on large, nationally-representative datasets. Across 100 years of research, teacher-assigned grades typically correlate about .5 with standardized measures of achievement" (Brookhart et al., 2016, p. 12). Since $r = .5$ is a moderate correlation coefficient, it simultaneously counters the argument that grades are a strong measure of academic skill and refutes the argument that grades are arbitrarily subjective measures of content knowledge (Linn, 1982). A correlation of $r = .5$ indicates 75% of the variation of grades assigned by teachers is attributable to something other than a trait comparable to the trait measured by the designated standardized test (Bowers, 2011; Brookhart et al., 2016). Thus, the previous research suggests that grades are multi-dimensional factors that consist of academic knowledge and cognitive achievement such as student self-efficacy, coping strategies, motivation, subject-specific interest, and work habits (Klapp Lekholm, 2011; Klapp Lekholm, & Cliffordson, 2008, 2009).

Summary

Chapter 2 contained the history of grading in the United States, including the rise of standards-based grading. Subsequent sections provided discourse on the reliability of grades, how grades are configured, how grades are used to predict educational outcomes, and the history of the ACT. The final section reviewed the relationship, as determined by previous studies, between standardized assessment scores and GPAs. Chapter 3 contains an explanation of the methodology used in this study, including the research design, selection of participants, measurement, data collection procedures, data analysis and hypothesis testing, and limitations.

Chapter 3

Methods

The purpose of this study was to examine the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs in grades 9 through 12 in high schools in three south-central Kansas school districts. An additional purpose of this study was to explore how the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs was impacted by gender, race, socioeconomic status, and high school community type. This chapter focuses on the methodology of the study including research design, selection of participants, measurement, data collection procedures, data analysis and hypothesis testing, and limitations.

Research Design

A quantitative correlational research design was used in this study, which utilized archived student data from three south-central Kansas school districts to analyze the relationship between ACT score and core GPA. Lunenburg and Irby (2008) stated that "correlational research…compares the degree to which the variables are related" (p. 35). Scores from two or more variables from the same sample are compared for the direction and strength of the relationship between them (Lunenburg & Irby, 2008). The variables of interest in this study were ACT score and core GPA. Additionally, gender, race, socioeconomic status, and community type were examined to determine if the demographics affected the relationship between ACT scores and core GPAs.

Selection of Participants

The population for the current study included students who took the ACT in grades 9 through 12 and graduated from a south-central Kansas high school. Students with ACT scores from 13 high schools in three districts comprised the population. Purposive sampling was used. Lunenburg and Irby (2008) indicated that purposive sampling "involves selecting a sample based on the researcher's experience or knowledge of the group to be sampled" (p. 175). The sample consisted of students who took the ACT during the 2015-2016 or 2016-2017 school year and graduated in the spring of 2017 from Sunflower Public Schools, Daisy Public Schools, or Marigold Public Schools. Of the students in the sample, 1,149 graduated from Sunflower Public Schools, 480 graduated from Daisy Public Schools, and 95 graduated from Marigold Public Schools.

Measurement

Six quantitative variables were measured in the current study: English ACT score, mathematics ACT score, reading ACT score, science ACT score, composite ACT score, and core GPA. Core GPA is a numeric average of all earned grades in the four major academic areas during grades 9 through 12: English, mathematics, history, and science (Garton et al., 2002). Since core GPA is an average of grades on student transcripts, core GPA is an accurate record of student grades in English, mathematics, history, and science despite student mobility or course selection. However, as noted in the literature review, grading practices vary substantially among teachers, schools, and districts, therefore it is nearly impossible to determine the reliability of core GPA.

The ACT is a multiple-choice exam which consists of four subject specific sub-tests: English, mathematics, reading, and science (ACT, 2014).

> The ACT English Test is a 75-item, 45-minute test that measures understanding of the conventions of standard written English (punctuation, grammar and usage, and sentence structure) and of rhetorical skills (strategy, organization, and style). The ACT Mathematics Test is a 60-item, 60-minute test that is designed to assess the mathematical reasoning skills that students across the United States have typically acquired in courses taken up to the beginning of Grade 12. The ACT Reading Test is a 40-item, 35-minute test that measures reading comprehension as a product of skill in referring and reasoning. The ACT Science Test is a 40-item, 35-minute test that measures the interpretation, analysis, evaluation, reasoning, and problem-solving skills required in the natural sciences. (ACT, 2014, pp. 5-6)

Each sub-test is scaled from 1 (low) to 36 (high) based on the number of correctly identified answers (ACT, 2014). The composite ACT score is the non-weighted average of the four sub-test scaled scores (ACT, 2014).

Technical measures are employed to ensure the ACT has the same statistical properties over time, since the test is constructed to measure the same content from version to version and from year to year (ACT, 2005). The ACT (2014) organization has conducted convergent and discriminant validity tests to evaluate the effectiveness of the ACT in measuring college readiness. There is a moderate (r = .66) to strong (r = .88) relationship between prior standardized test scores and composite ACT scores (ACT, 2014). "The strength of the relationships provides evidence for the criterion validity of the [ACT]" (ACT, 2014, p. 87).

According to Lunenburg and Irby (2008), reliability is the degree to which an instrument consistently measures a certain characteristic. "Reliability coefficients are estimates of the constancy of test scores. They typically range from zero to one, with values near one indicating greater consistency and those near zero indicating little or no consistency" (ACT, 2014, p. 51). The reliability of the composite ACT score has ranged from .96 to .97, between 1995 and 2012, as measured by internal consistency and test-retest stability reliability tests (ACT, 2014). While no instrument can claim complete validity and reliability, the evidence above provides strong indications for the validity and reliability of the ACT (ACT, 2014).

Four categorical variables were used in this study: gender, race, socioeconomic status, and community type. Student gender (male or female) was indicated on each student's personal information form. Student race was indicated on each student's school record. The following race categories were included in this study: Black, White, Hispanic, Asian, and other. For the purpose of the data analysis students categorized as Native American, Alaskan Native, Native Hawaiian, other Pacific Islander, or Multiracial were included in the other category since those racial groups were less prevalent in the suburban and rural districts. Student socioeconomic status (students who qualified for free or reduced lunch or students who qualified for full price lunch) was determined by federal guidelines for parent or guardian income as part of the district database. Community type was determined by geographic location and U.S. Census Bureau categorization of the school district from which a student graduated. Students from Marigold Public Schools were considered part of a rural community since Marigold, Kansas is between five and 25 miles from an urbanized area (Geverdt, 2015). Daisy Public Schools' students were labeled suburban because Daisy, Kansas has a population well below 100,000 and is outside of a principal city (Geverdt, 2015). With a population of well over 250,000 within the principal city of Sunflower, Kansas, students from Sunflower Public Schools were categorized as urban (Geverdt, 2015).

Data Collection Procedures

Prior to collecting data, permission to conduct the study in Sunflower Public Schools was acquired by completing a research request form on November 16, 2016 (Appendix A). The research request form was sent to the Executive Director of Innovation and Evaluation, who examined the request and approved the research with the Sunflower Public Schools' Innovation and Learning Committee on December 5, 2016 (Appendix B). After obtaining approval from Sunflower Public Schools, Daisy Public Schools (Appendix C) and Marigold Public Schools (Appendix D) were contacted via e-mail on December 9, 2016. On December 12, 2016, Daisy Public Schools' superintendent agreed to supply data for the study contingent on receiving a copy of the results (Appendix E). Similarly, on December 15, 2016, Marigold Public Schools' superintendent agreed to provide district data for the study as long as the district office received three spiral bound copies of the finalized report (Appendix F). Permission to conduct the study was also requested from the Baker University Institutional Review Board (IRB) on April 3, 2018 (Appendix G). The IRB committee approved the IRB form on April 4, 2018 (Appendix H).

The following data was requested, via e-mail, in an Excel spreadsheet: English ACT score, mathematics ACT score, reading ACT score, science ACT score, composite ACT score, core GPA, gender, race, and socioeconomic status. If a student had more than one ACT score on file, the districts were instructed to only include the most recent score. On April 6, 2018, the Executive Director of Innovation and Evaluation provided an Excel spreadsheet for Sunflower Public Schools. The Excel spreadsheet containing Daisy Public Schools' data was provided by the Assistant Superintendent for Teaching and Learning on April 9, 2018. On April 11, 2018, the Assistant Superintendent in Marigold Public Schools provided an Excel spreadsheet. The three Excel spreadsheets were merged, and a community type column was added based on the district's classification as urban, suburban, or rural. Student names were not included in any spreadsheet. The data was verified and stored on a password protected external hard drive.

Data Analysis and Hypothesis Testing

Archived quantitative data were used in this study. The data supplied by the districts were organized into a Microsoft Excel worksheet and then imported into IBM ® SPSS ® Statistics Faculty Pack 24 for Mac. The data were used to answer the two research questions. Twenty-five hypotheses were tested for a statistically significant correlation between core GPA and ACT score, and differences between correlations based on gender, race, socioeconomic status, and community type.

RQ1. To what extent is there a relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs for students in grades 9 through 12 in three south-central Kansas districts?

***H1*.** There is a relationship between English ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts.

A Pearson product moment correlation coefficient was calculated to determine the strength and direction of the relationship between English ACT scores and core GPAs.

An independent-samples *t* test was conducted to test for the statistical significance of the Pearson correlation coefficient. The level of significance was set at .05.

H2. There is a relationship between mathematics ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts.

A Pearson product moment correlation coefficient was calculated to determine the strength and direction of the relationship between mathematics ACT scores and core GPAs. An independent-samples *t* test was conducted to test for the statistical significance of the Pearson correlation coefficient. The level of significance was set at .05.

H3. There is a relationship between reading ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts.

A Pearson product moment correlation coefficient was calculated to determine the strength and direction of the relationship between reading ACT scores and core GPAs. An independent-samples *t* test was conducted to test for the statistical significance of the Pearson correlation coefficient. The level of significance was set at .05.

H4. There is a relationship between science ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts.

A Pearson product moment correlation coefficient was calculated to determine the strength and direction of the relationship between science ACT scores and core GPAs. An independent-samples *t* test was conducted to test for the statistical significance of the Pearson correlation coefficient. The level of significance was set at .05.

H5. There is a relationship between composite ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts.

A Pearson product moment correlation coefficient was calculated to determine the strength and direction of the relationship between composite ACT scores and core GPAs. An independent-samples *t* test was conducted to test for the statistical significance of the Pearson correlation coefficient. The level of significance was set at .05.

RQ2. To what extent is the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs for students in grades 9 through 12 in three south-central Kansas districts affected by gender, race, socioeconomic status, and community type?

H6. There is a difference in the relationship between English ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on gender.

Prior to testing H6, the sample data was disaggregated by gender (male, female). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between English ACT scores and core GPAs for males and females. A Fisher's *z* test for two correlations was conducted to determine if the difference in the relationship between English ACT scores and core GPAs for female students and male students was statistically significant. The sample correlation for female students was compared to the sample correlation for male students. The level of significance was set at .05.

H7. There is a difference in the relationship between mathematics ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on gender.

Prior to testing H7, the sample data was disaggregated by gender (male, female). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between mathematics ACT scores and core GPAs for males and females. A Fisher's *z* test for two correlations was conducted to determine if the difference in the relationship between mathematics ACT scores and core GPAs for female students and male students was statistically significant. The sample correlation for female students was compared to the sample correlation for male students. The level of significance was set at .05.

***H8*.** There is a difference in the relationship between reading ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on gender.

Prior to testing H8, the sample data was disaggregated by gender (male, female). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between reading ACT scores and core GPAs for males and females. A Fisher's *z* test for two correlations was conducted to determine if the difference in the relationship between reading ACT scores and core GPAs for female students and male students was statistically significant. The sample correlation for female students was compared to the sample correlation for male students. The level of significance was set at .05.

***H9*.** There is a difference in the relationship between science ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on gender.

Prior to testing H9, the sample data was disaggregated by gender (male, female). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between science ACT scores and core GPAs for males and females. A Fisher's *z* test for two correlations was conducted to determine if the difference in the relationship between science ACT scores and core GPAs for female students and male students was statistically significant. The sample correlation for female students was compared to the sample correlation for male students. The level of significance was set at .05.

H10. There is a difference in the relationship between composite ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on gender.

Prior to testing H10, the sample data was disaggregated by gender (male, female). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between composite ACT scores and core GPAs for males and females. A Fisher's *z* test for two correlations was conducted to determine if the difference in the relationship between composite ACT scores and core GPAs for female students and male students was statistically significant. The sample correlation for female students was compared to the sample correlation for male students. The level of significance was set at .05.

H11. There is a difference in the relationship between English ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among different races.

Prior to testing H11, the sample data was disaggregated by race (white, black, Hispanic, Asian, other). Five Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between English ACT scores and core GPAs for white, black, Hispanic, Asian, and other students (Native American, Alaskan Native, Native Hawaiian, other Pacific Islander, or Multiracial). Ten Fisher's *z* tests for two correlations were conducted to determine if the differences in the relationship between English ACT scores and core GPAs for white, black, Hispanic, Asian, and the other students were statistically significant. The first four Fisher's *z* tests for two correlations compared the sample correlation for white students to the sample correlation for black students, Hispanic students, Asian students, and other students respectively. The next three Fisher's *z* tests for two correlations compared the sample correlation for black students to the sample correlation for Hispanic students, Asian students, and other students respectively. The next two Fisher's *z* tests for two correlations compared the sample correlation for Hispanic students to the sample correlation for Asian students and other students respectively. Finally, the last Fisher's *z* test for two correlations compared the sample correlation for Asian students to the sample correlation for other students. The level of significance was set at .05.

H12. There is a difference in the relationship between mathematics ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among different races.

Prior to testing H12, the sample data was disaggregated by race (white, black, Hispanic, Asian, other). Five Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between mathematics ACT scores and core GPAs for white, black, Hispanic, Asian, and other students (Native American, Alaskan Native, Native Hawaiian, other Pacific Islander, or Multiracial). Ten Fisher's *z* tests for two correlations were conducted to determine if the differences in the relationship between mathematics ACT scores and core GPAs for white, black, Hispanic, Asian, and the other students were statistically significant. The first four Fisher's *z* tests for two correlations compared the sample correlation for white students to the sample correlation for black students, Hispanic students, Asian students, and other students respectively. The next three Fisher's *z* tests for two correlations compared the sample correlation for black students to the sample correlation for Hispanic students, Asian students, and other students respectively. The next two Fisher's *z* tests for two correlations compared the sample correlation for Hispanic students to the sample correlation for Asian students and other students respectively. Finally, the last Fisher's *z* test for two correlations compared the sample correlation for Asian students to the sample correlation for other students. The level of significance was set at .05.

***H13*.** There is a difference in the relationship between reading ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among different races.

Prior to testing H13, the sample data was disaggregated by race (white, black, Hispanic, Asian, other). Five Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between reading ACT scores and core GPAs for white, black, Hispanic, Asian, and other students (Native American, Alaskan Native, Native Hawaiian, other Pacific Islander, or Multiracial). Ten Fisher's *z* tests for two correlations were conducted to determine if the differences in the relationship between reading ACT scores and core GPAs for white, black, Hispanic, Asian, and the other students were statistically significant. The first four Fisher's *z* tests for two correlations compared the sample correlation for white students to the sample correlation for black students, Hispanic students, Asian students, and other students respectively. The next three Fisher's *z* tests for two correlations compared the sample correlation for black students to the sample correlation for Hispanic students, Asian students, and other students respectively. The next two Fisher's *z* tests for two correlations compared the sample correlation for Hispanic students to the sample correlation for Asian students and other students respectively. Finally, the last Fisher's *z* test for two correlations compared the sample correlation for Asian students to the sample correlation for other students. The level of significance was set at .05.

H14. There is a difference in the relationship between science ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among different races.

Prior to testing H14, the sample data was disaggregated by race (white, black, Hispanic, Asian, other). Five Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between science ACT scores and core GPAs for white, black, Hispanic, Asian, and other students (Native American, Alaskan Native, Native Hawaiian, other Pacific Islander, or Multiracial). Ten Fisher's *z* tests for two correlations were conducted to determine if the differences in the relationship between science ACT scores and core GPAs for white, black, Hispanic, Asian, and the other students were statistically significant. The first four Fisher's *z* tests for two correlations compared the sample correlation for white students to the sample correlation for black students, Hispanic students, Asian students, and other students respectively. The next three Fisher's *z* tests for two correlations compared the sample correlation for black students to the sample correlation for Hispanic students, Asian students, and other students respectively. The next two Fisher's *z* tests for two correlations compared the sample correlation for Hispanic students to the sample correlation for Asian students and other students respectively. Finally, the last Fisher's *z* test for two correlations compared the sample correlation for Asian students to the sample correlation for other students. The level of significance was set at .05.

H15. There is a difference in the relationship between composite ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among different races.

Prior to testing H15, the sample data was disaggregated by race (white, black, Hispanic, Asian, other). Five Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between composite ACT scores and core GPAs for white, black, Hispanic, Asian, and other students (Native American, Alaskan Native, Native Hawaiian, other Pacific Islander, or Multiracial). Ten Fisher's *z* tests for two correlations were conducted to determine if the differences in the relationship between composite ACT scores and core GPAs for white, black, Hispanic, Asian, and the other students were statistically significant. The first four Fisher's *z* tests for two correlations compared the sample correlation for white students to the sample correlation for black students, Hispanic students, Asian students, and other students respectively. The next three Fisher's *z* tests for two correlations compared the sample correlation for black students to the sample correlation for Hispanic students, Asian students, and other students respectively. The next two Fisher's *z* tests for two correlations compared the sample correlation for Hispanic students to the sample correlation for Asian students and other students respectively. Finally, the last Fisher's *z* test for two correlations compared the sample correlation for Asian students to the sample correlation for other students. The level of significance was set at .05.

H16. There is a difference in the relationship between English ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on SES level.

Prior to testing H16, the sample data was disaggregated by SES level (students who qualified for free or reduced lunch, students who qualified for full price lunch). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between English ACT scores and core GPAs. A Fisher's *z* test for two correlations was conducted to determine if the difference in relationship between English ACT scores and core GPAs for students who qualified for free or reduced lunch and students who qualified for full price lunch was statistically significant. The sample correlation for students who qualified for free or reduced lunch was compared to the sample correlation for students who qualified for full price lunch. The level of significance was set at .05.

H17. There is a difference in the relationship between mathematics ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on SES level.

Prior to testing H17, the sample data was disaggregated by SES level (students who qualified for free or reduced lunch, students who qualified for full price lunch). Two

Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between mathematics ACT scores and core GPAs. A Fisher's *z* test for two correlations was conducted to determine if the difference in relationship between mathematics ACT scores and core GPAs for students who qualified for free or reduced lunch and students who qualified for full price lunch was statistically significant. The sample correlation for students who qualified for free or reduced lunch was compared to the sample correlation for students who qualified for full price lunch. The level of significance was set at .05.

H18. There is a difference in the relationship between reading ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on SES level.

Prior to testing H18, the sample data was disaggregated by SES level (students who qualified for free or reduced lunch, students who qualified for full price lunch). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between reading ACT scores and core GPAs. A Fisher's *z* test for two correlations was conducted to determine if the difference in relationship between reading ACT scores and core GPAs for students who qualified for free or reduced lunch and students who qualified for full price lunch was statistically significant. The sample correlation for students who qualified for free or reduced lunch was compared to the sample correlation for students who qualified for full price lunch. The level of significance was set at .05.

***H19*.** There is a difference in the relationship between science ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on SES level.

Prior to testing H19, the sample data was disaggregated by SES level (students who qualified for free or reduced lunch, students who qualified for full price lunch). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between science ACT scores and core GPAs. A Fisher's *z* test for two correlations was conducted to determine if the difference in relationship between science ACT scores and core GPAs for students who qualified for free or reduced lunch and students who qualified for full price lunch was statistically significant. The sample correlation for students who qualified for free or reduced lunch was compared to the sample correlation for students who qualified for full price lunch. The level of significance was set at .05.

***H20*.** There is a difference in the relationship between composite ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on SES level.

Prior to testing H20, the sample data was disaggregated by SES level (students who qualified for free or reduced lunch, students who qualified for full price lunch). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between composite ACT scores and core GPAs. A Fisher's *z* test for two correlations was conducted to determine if the difference in relationship between composite ACT scores and core GPAs for students who qualified for free or reduced lunch and students who qualified for full price lunch was statistically significant. The sample correlation for students who qualified for free or reduced lunch was compared to the sample correlation for students who qualified for full price lunch. The level of significance was set at .05.

H21. There is a difference in the relationship between English ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among community types.

Prior to testing H21, the sample data was disaggregated by community type (urban, suburban, and rural). Three Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between English ACT scores and core GPAs for urban, suburban, and rural community type. Three Fisher's *z* tests for two correlations were conducted to determine the difference in relationship between English ACT scores and core GPAs for urban students, suburban students, and rural students. The first Fisher's *z* test for two correlations compared the sample correlation for urban students to the sample correlation for suburban students. The second Fisher's *z* test for two correlations compared the sample correlation for urban students to the sample correlation for rural students. The third Fisher's *z* test for two correlations compared the sample correlation for suburban students to the sample correlation for rural students. The level of significance was set at .05.

H22. There is a difference in the relationship between mathematics ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among community types.

Prior to testing H22, the sample data was disaggregated by community type (urban, suburban, and rural). Three Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between mathematics ACT scores and core GPAs for urban, suburban, and rural community type. Three Fisher's *z* tests for two correlations were conducted to determine the difference in relationship between mathematics ACT scores and core GPAs for urban students, suburban students, and rural students. The first Fisher's *z* test for two correlations compared the sample correlation for urban students to the sample correlation for suburban students. The second Fisher's *z* test for two correlations compared the sample correlation for urban students to the sample correlation for rural students. The third Fisher's *z* test for two correlations compared the sample correlation for suburban students to the sample correlation for rural students. The level of significance was set at .05.

***H23*.** There is a difference in the relationship between reading ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among community types.

Prior to testing H23, the sample data was disaggregated by community type (urban, suburban, and rural). Three Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between reading ACT scores and core GPAs for urban, suburban, and rural community type. Three Fisher's *z* tests for two correlations were conducted to determine the difference in relationship between reading ACT scores and core GPAs for urban students, suburban students, and rural students. The first Fisher's *z* test for two correlations compared the sample correlation for urban students to the sample correlation for suburban students. The second Fisher's *z* test for two correlations compared the sample correlation for urban students to the sample correlation for rural students. The third Fisher's *z* test for two correlations compared the sample correlation for suburban students to the sample correlation for rural students. The level of significance was set at .05.

H24. There is a difference in the relationship between science ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among community types.

Prior to testing H24, the sample data was disaggregated by community type (urban, suburban, and rural). Three Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between science ACT scores and core GPAs for urban, suburban, and rural community type. Three Fisher's *z* tests for two correlations were conducted to determine the difference in relationship between science ACT scores and core GPAs for urban students, suburban students, and rural students. The first Fisher's *z* test for two correlations compared the sample correlation for urban students to the sample correlation for suburban students. The second Fisher's *z* test for two correlations compared the sample correlation for urban students to the sample correlation for rural students. The third Fisher's *z* test for two correlations compared the sample correlation for suburban students to the sample correlation for rural students. The level of significance was set at .05.

H25. There is a difference in the relationship between composite ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among community types.

Prior to testing H25, the sample data was disaggregated by community type (urban, suburban, and rural). Three Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between composite ACT scores and core GPAs for urban, suburban, and rural community type. Three Fisher's *z* tests for two correlations were conducted to determine the difference in relationship between composite ACT scores and core GPAs for urban students, suburban students, and rural students. The first Fisher's *z* test for two correlations compared the sample correlation for urban students to the sample correlation for suburban students. The second Fisher's *z* test for two correlations compared the sample correlation for urban students to the sample correlation for rural students. The third Fisher's *z* test for two correlations compared the sample correlation for suburban students to the sample correlation for rural students. The level of significance was set at .05.

Limitations

"Limitations are factors that may have an effect on the interpretation of the findings or the generalizability of the results" (Lunenburg & Irby, 2008, p. 131). Even though researchers cannot control limitations, providing the reader with limitations can help avoid misinterpretation of the findings (Lunenburg & Irby, 2008). The limitations for this study included:

1. The sample was limited to the students who chose to take the ACT. Students who only took the SAT or did not take a college entrance exam were excluded.
2. The study was limited to the data supplied from each district, so accuracy and completeness of the data cannot be independently verified.
3. Test preparation and instructional strategies vary between high schools within a single district and between districts.
4. Student ACT scores may be affected by amount of sleep the night before, health at the time of the test, testing environment, and intrinsic motivation.

Summary

Chapter 3 described the methodology used in this study. The research design, selection of participants, measurement, data collection procedures, data analysis and hypothesis testing, and limitations were all included in this chapter. Chapter 4 includes an explanation of the descriptive statistics, hypothesis testing results, and a summary.

Chapter 4

Results

The major purpose of this study was to examine the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs in grades 9 through 12 in high schools in three south-central Kansas districts. A secondary purpose of this study was to explore the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs as impacted by gender, race, socioeconomic status, and high school community type. This chapter focuses on descriptive statistics for ACT scores and core GPAs disaggregated by gender, race, socioeconomic status, and community type and the results of the hypothesis testing.

Descriptive Statistics

Research question one focuses on the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs and involves five hypotheses. The mean GPA (M = 3.12, SD = 0.61) indicates the average student in the current study earns all Bs and one A each semester. See Table 4 for the means and standard deviations for the ACT scores used in the analysis of research question one. The mean for reading ACT scores (M = 21.09) was the highest of all the ACT scores while the mean for English ACT scores (M = 19.43) was the lowest.

Table 4

ACT Scores: Descriptive Statistics

ACT Test	*M*	*SD*	*N*
English	19.43	6.10	1724
Mathematics	20.34	4.70	1724
Reading	21.09	3.06	1724
Science	20.56	4.88	1724
Composite	20.48	4.95	1724

Research question two focuses on how gender, race, socioeconomic status, and community type affect the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs and involves twenty hypotheses. See Table 5 for the means and standard deviations for the ACT scores disaggregated by gender. The mean for male students was higher than the mean for female students for mathematics, science, and composite ACT scores.

Table 5

ACT Scores Disaggregated by Gender

ACT Test	Gender	*M*	*SD*	*N*
English	Female	19.52	6.08	920
	Male	19.33	6.13	804
Mathematics	Female	19.97	4.57	920
	Male	20.78	4.98	804
Reading	Female	21.14	6.01	920
	Male	21.03	6.06	804
Science	Female	20.38	4.66	920
	Male	20.76	5.11	804
Composite	Female	20.36	4.85	920
	Male	20.61	5.06	804

See Table 6 for the means and standard deviations for the core GPAs disaggregated by gender. Despite lower means in three of the five ACT scores, female students have a higher mean core GPA (M = 3.21) than male students (M = 3.01).

Table 6

Core GPAs Disaggregated by Gender

Gender	*M*	*SD*	*N*
Female	3.21	0.58	920
Male	3.01	0.63	804

See Table 7 for the means and standard deviations for the ACT scores disaggregated by race. The mean for white students was higher than the mean for any other race with the exception of Asian students on the mathematics section of the ACT.

Table 7

ACT Scores Disaggregated by Race

ACT Test	Race	*M*	*SD*	*N*
English	White	21.25	6.07	889
	Black	15.85	5.21	224
	Hispanic	16.62	4.86	313
	Asian	20.04	5.75	133
	Other	19.35	5.37	165
Mathematics	White	21.54	4.89	889
	Black	17.66	3.56	224
	Hispanic	18.50	3.92	313
	Asian	21.67	4.87	133
	Other	19.96	4.42	165
Reading	White	22.64	6.00	889
	Black	17.97	5.02	224
	Hispanic	18.69	5.07	313
	Asian	21.35	6.35	133
	Other	21.31	5.82	165
Science	White	21.94	4.82	889
	Black	17.55	4.11	224
	Hispanic	18.76	4.16	313
	Asian	21.17	4.97	133
	Other	20.08	4.40	165
Composite	White	21.98	4.88	889
	Black	17.36	3.98	224
	Hispanic	18.29	3.99	313
	Asian	21.16	4.94	133
	Other	20.30	4.59	165

See Table 8 for the means and standard deviations for the core GPAs disaggregated by race. Despite the highest means in four of the five ACT scores, white students do not have the highest core GPA (*M* = 3.24), Asian students have the highest core GPA (*M* = 3.43).

Table 8

Core GPAs Disaggregated by Race

Race	*M*	*SD*	*N*
White	3.24	0.58	889
Black	2.83	0.62	224
Hispanic	2.86	0.60	313
Asian	3.43	0.49	133
Other	3.08	0.58	165

See Table 9 for the means and standard deviations for the ACT scores disaggregated by socioeconomic status. The means for students who qualified for free or reduced lunch was lower than the means for students who qualified for full price lunch in all four ACT subject tests and composite ACT scores.

Table 9

ACT Scores Disaggregated by Socioeconomic Status

ACT Test	SES	*M*	*SD*	*N*
English	Free or Reduced	17.17	5.40	709
	Full Price	21.01	6.07	1015
Mathematics	Free or Reduced	18.72	4.13	709
	Full Price	21.48	4.88	1015
Reading	Free or Reduced	19.00	5.37	709
	Full Price	22.55	6.04	1015
Science	Free or Reduced	18.89	4.56	709
	Full Price	21.72	4.76	1015
Composite	Free or Reduced	18.55	4.34	709
	Full Price	21.82	4.90	1015

See Table 10 for the means and standard deviations for the core GPAs disaggregated by socioeconomic status. The mean core GPA for students who qualified for free or reduced lunch (M = 2.92) is lower than the mean core GPA for students who qualified for full price lunch (M = 3.29).

Table 10

Core GPAs Disaggregated by Socioeconomic Status

SES	*M*	*SD*	*N*
Free or Reduced	2.92	0.60	709
Full Price	3.29	0.58	1015

See Table 11 for the means and standard deviations for the ACT scores disaggregated by community type. The means for students in an urban high school were lower than the means for students in a suburban or a rural high school. Students in a suburban high school had the highest mean ACT score in all four ACT subject tests and composite ACT scores.

Table 11

ACT Scores Disaggregated by Community Type

ACT Test	Community Type	*M*	*SD*	*N*
English	Urban	18.33	5.86	1149
	Suburban	21.88	6.02	480
	Rural	20.41	5.62	95
Mathematics	Urban	19.59	4.56	1149
	Suburban	22.22	4.92	480
	Rural	20.00	4.12	95
Reading	Urban	20.04	5.82	1149
	Suburban	23.38	5.92	480
	Rural	22.17	5.75	95
Science	Urban	19.68	4.73	1149
	Suburban	22.45	4.84	480
	Rural	21.51	3.86	95
Composite	Urban	19.53	4.73	1149
	Suburban	22.62	4.90	480
	Rural	21.11	4.21	95

See Table 12 for the means and standard deviations for the core GPAs disaggregated by community type. The mean core GPA for students in an urban high school (M = 3.02) is lower than the mean core GPA for students in a suburban high school (M = 3.33) and students in a rural high school (M = 3.26).

Table 12

Core GPAs Disaggregated by Community Type

Community Type	*M*	*SD*	*N*
Urban	3.02	0.60	1149
Suburban	3.33	0.58	480
Rural	3.26	0.64	95

Hypothesis Testing

The first five analyses and results presented below are the outcomes of the hypothesis testing which addressed research question one for English, mathematics, reading, science, and composite ACT scores. The correlations between the ACT scores and core GPAs were then re-analyzed when disaggregated by gender, race, socioeconomic status, and community type. The final 20 analyses and results are the outcomes of the hypothesis testing which addressed research question two.

RQ1. To what extent is there a relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs for students in grades 9 through 12 in three south-central Kansas districts?

***H1*.** There is a relationship between English ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts.

A Pearson product moment correlation coefficient was calculated to determine the strength and direction of the relationship between English ACT scores and core GPAs. An independent-samples *t* test was conducted to test for the statistical significance of the Pearson correlation coefficient. The level of significance was set at .05. The correlation coefficient (r = .541) provided evidence for a moderate positive relationship between English ACT scores and core GPAs. The results of the independent-samples *t* test indicated a statistically significant relationship between English ACT scores and core GPAs, $df = 1722$, $p = .000$.

H2. There is a relationship between mathematics ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts.

A Pearson product moment correlation coefficient was calculated to determine the strength and direction of the relationship between mathematics ACT scores and core GPAs. An independent-samples t test was conducted to test for the statistical significance of the Pearson correlation coefficient. The level of significance was set at .05. The correlation coefficient ($r = .578$) provided evidence for a moderate positive relationship between mathematics ACT scores and core GPAs. The results of the independent-samples t test indicated a statistically significant relationship between mathematics ACT scores and core GPAs, $df = 1722$, $p = .000$.

H3. There is a relationship between reading ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts.

A Pearson product moment correlation coefficient was calculated to determine the strength and direction of the relationship between reading ACT scores and core GPAs. An independent-samples t test was conducted to test for the statistical significance of the Pearson correlation coefficient. The level of significance was set at .05. The correlation coefficient ($r = .478$) provided evidence for a moderate positive relationship between reading ACT scores and core GPAs. The results of the independent-samples t test indicated a statistically significant relationship between reading ACT scores and core GPAs, $df = 1722$, $p = .000$.

H4. There is a relationship between science ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts.

A Pearson product moment correlation coefficient was calculated to determine the strength and direction of the relationship between science ACT scores and core GPAs. An independent-samples *t* test was conducted to test for the statistical significance of the Pearson correlation coefficient. The level of significance was set at .05. The correlation coefficient (r = .516) provided evidence for a moderate positive relationship between science ACT scores and core GPAs. The results of the independent-samples *t* test indicated a statistically significant relationship between science ACT scores and core GPAs, df = 1722, p = .000.

H5. There is a relationship between composite ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts.

A Pearson product moment correlation coefficient was calculated to determine the strength and direction of the relationship between composite ACT scores and core GPAs. An independent-samples *t* test was conducted to test for the statistical significance of the Pearson correlation coefficient. The level of significance was set at .05. The correlation coefficient (r = .584) provided evidence for a moderate positive relationship between composite ACT scores and core GPAs. The results of the independent-samples *t* test indicated a statistically significant relationship between composite ACT scores and core GPAs, df = 1722, p = .000.

RQ2. To what extent is the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs for students in grades 9 through 12 in three south-central Kansas districts affected by gender, race, socioeconomic status, and community type?

***H6*.** There is a difference in the relationship between English ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on gender.

Prior to testing H6, the sample data was disaggregated by gender (male, female). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between English ACT scores and core GPAs for males and females. The level of significance was set at .05. The correlation coefficient (r = .528) provided evidence for a moderate positive relationship between English ACT scores and core GPAs for female students. The results of the independent-samples *t* test indicated a statistically significant relationship between English ACT scores and core GPAs for female students, df = 918, p = .000. The correlation coefficient (r = .566) provided evidence for a moderate positive relationship between English ACT scores and core GPAs for male students. The results of the independent-samples *t* test indicated a statistically significant relationship between English ACT scores and core GPAs for male students, df = 802, p = .000.

A Fisher's *z* test for two correlations was conducted to determine if the difference in the relationship between English ACT scores and core GPAs for female students and male students was statistically significant. The sample correlation for female students was compared to the sample correlation for male students. The level of significance was set at .05. The results of the Fisher's *z* test for two correlations indicated no difference between the two values, z = -1.12, p = .263. The correlation for female students

(r = .528) was not different from the correlation for male students (r = .566). The correlation between English ACT scores and core GPAs was not affected by gender.

H7. There is a difference in the relationship between mathematics ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on gender.

Prior to testing H7, the sample data was disaggregated by gender (male, female). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between mathematics ACT scores and core GPAs for males and females. The level of significance was set at .05. The correlation coefficient (r = .588) provided evidence for a moderate positive relationship between mathematics ACT scores and core GPAs for female students. The results of the independent-samples t test indicated a statistically significant relationship between mathematics ACT scores and core GPAs for female students, df = 918, p = .000. The correlation coefficient (r = .616) provided evidence for a moderately strong positive relationship between mathematics ACT scores and core GPAs for male students. The results of the independent-samples t test indicated a statistically significant relationship between mathematics ACT scores and core GPAs for male students, df = 802, p = .000.

A Fisher's z test for two correlations was conducted to determine if the difference in the relationship between mathematics ACT scores and core GPAs for female students and male students was statistically significant. The sample correlation for female students was compared to the sample correlation for male students. The level of significance was set at .05. The results of the Fisher's z test for two correlations indicated no difference between the two values, z = -0.91, p = .363. The correlation for female students (r = .588) was not different from the correlation for male students (r = .616). The correlation between mathematics ACT scores and core GPAs was not affected by gender.

H8. There is a difference in the relationship between reading ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on gender.

Prior to testing H8, the sample data was disaggregated by gender (male, female). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between reading ACT scores and core GPAs for males and females. The level of significance was set at .05. The correlation coefficient (r = .470) provided evidence for a moderate positive relationship between reading ACT scores and core GPAs for female students. The results of the independent-samples t test indicated a statistically significant relationship between reading ACT scores and core GPAs for female students, df = 918, p = .000. The correlation coefficient (r = .498) provided evidence for a moderate positive relationship between reading ACT scores and core GPAs for male students. The results of the independent-samples t test indicated a statistically significant relationship between reading ACT scores and core GPAs for male students, df = 802, p = .000.

A Fisher's z test for two correlations was conducted to determine if the difference in the relationship between reading ACT scores and core GPAs for female students and male students was statistically significant. The sample correlation for female students was compared to the sample correlation for male students. The level of significance was set at .05. The results of the Fisher's *z* test for two correlations indicated no difference between the two values, $z = -0.76$, $p = .447$. The correlation for female students ($r = .470$) was not different from the correlation for male students ($r = .498$). The correlation between reading ACT scores and core GPAs was not affected by gender.

***H9*.** There is a difference in the relationship between science ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on gender.

Prior to testing H9, the sample data was disaggregated by gender (male, female). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between science ACT scores and core GPAs for males and females. The level of significance was set at .05. The correlation coefficient ($r = .509$) provided evidence for a moderate positive relationship between science ACT scores and core GPAs for female students. The results of the independent-samples *t* test indicated a statistically significant relationship between science ACT scores and core GPAs for female students, $df = 918$, $p = .000$. The correlation coefficient ($r = .552$) provided evidence for a moderate positive relationship between science ACT scores and core GPAs for male students. The results of the independent-samples *t* test indicated a statistically significant relationship between science ACT scores and core GPAs for male students, $df = 802$, $p = .000$.

A Fisher's *z* test for two correlations was conducted to determine if the difference in the relationship between science ACT scores and core GPAs for female students and male students was statistically significant. The sample correlation for female students was compared to the sample correlation for male students. The level of significance was set at .05. The results of the Fisher's *z* test for two correlations indicated no difference between the two values, $z = -1.24$, $p = .215$. The correlation for female students ($r = .509$) was not different from the correlation for male students ($r = .552$). The correlation between science ACT scores and core GPAs was not affected by gender.

H10. There is a difference in the relationship between composite ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on gender.

Prior to testing H10, the sample data was disaggregated by gender (male, female). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between composite ACT scores and core GPAs for males and females. The level of significance was set at .05. The correlation coefficient ($r = .578$) provided evidence for a moderate positive relationship between composite ACT scores and core GPAs for female students. The results of the independent-samples *t* test indicated a statistically significant relationship between composite ACT scores and core GPAs for female students, $df = 918$, $p = .000$. The correlation coefficient ($r = .616$) provided evidence for a moderately strong positive relationship between composite ACT score and core GPAs for male students. The results of the independent-samples *t* test indicated a statistically significant relationship between composite ACT scores and core GPAs for male students, $df = 802$, $p = .000$.

A Fisher's *z* test for two correlations was conducted to determine if the difference in the relationship between composite ACT scores and core GPAs for female students and male students was statistically significant. The sample correlation for female students was compared to the sample correlation for male students. The level of significance was set at .05. The results of the Fisher's *z* test for two correlations indicated no difference between the two values, $z = -1.22$, $p = .223$. The correlation for female students ($r = .578$) was not different from the correlation for male students ($r = .616$). The correlation between composite ACT scores and core GPAs was not affected by gender.

H11. There is a difference in the relationship between English ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among different races.

Prior to testing H11, the sample data was disaggregated by race (white, black, Hispanic, Asian, other). Five Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between English ACT scores and core GPAs for white, black, Hispanic, Asian, and other students (Native American, Alaskan Native, Native Hawaiian, other Pacific Islander, or Multiracial). The level of significance was set at .05. Ten Fisher's *z* tests for two correlations were conducted to determine if the differences in the relationship between English ACT scores and core GPAs for white, black, Hispanic, Asian, and the other students were statistically significant. The level of significance was set at .05 for each Fisher's *z* test for two correlations. See Table 13 for the results of the calculation of the Pearson product moment correlation coefficients and the Fisher's *z* tests for two correlations for English ACT scores and core GPAs disaggregated by race.

Table 13

Results for English ACT Scores and Core GPAs Disaggregated by Race

Subgroup 1	*r(df)*	Subgroup 2	*r(df)*	*z*	*p*
White	.522*** (887)	Black	.492*** (222)	0.54	.589
White	.522*** (887)	Hispanic	.413*** (311)	2.12	.034
White	.522*** (887)	Asian	.430*** (131)	1.27	.240
White	.522*** (887)	Other	.538*** (163)	-0.26	.795
Black	.492*** (222)	Hispanic	.413*** (311)	1.13	.259
Black	.492*** (222)	Asian	.430*** (131)	0.71	.478
Black	.492*** (222)	Other	.538*** (163)	-0.61	.542
Hispanic	.413*** (311)	Asian	.430*** (131)	-0.20	.842
Hispanic	.413*** (311)	Other	.538*** (163)	-1.61	.095
Asian	.430*** (131)	Other	.538*** (163)	-1.20	.230

Note. *p < .05, **p < .01, ***p <.001

The correlation coefficients provided evidence for a moderate positive relationship between English ACT scores and core GPAs for each race subgroup. The results of the independent-samples *t* test indicated a statistically significant relationship between English ACT scores and core GPAs for each race subgroup. The results of the Fisher's *z* tests for two correlations indicated no difference between the two values when comparing white students to black students, white students to Asian students, white students to students in the other race category, black students to Hispanic students, black students to Asian students, black students to students in the other race category, Hispanic students to Asian students, Hispanic students to students in the other race category, and Asian students to students in the other race category. The results of the Fisher's *z* test for two correlations indicated a statistically significant difference between the two values when comparing white students to Hispanic students. The correlation for white students was stronger than the correlation for Hispanic students. The correlation between English ACT scores and core GPAs was affected by race.

H12. There is a difference in the relationship between mathematics ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among different races.

Prior to testing H12, the sample data was disaggregated by race (white, black, Hispanic, Asian, other). Five Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between mathematics ACT scores and core GPAs for white, black, Hispanic, Asian, and other students (Native American, Alaskan Native, Native Hawaiian, other Pacific Islander, or Multiracial). The level of significance was set at .05. Ten Fisher's *z* tests for two correlations were conducted to determine if the differences in the relationship between mathematics ACT scores and core GPAs for white, black, Hispanic, Asian, and the other students were statistically significant. The level of significance was set at .05 for each Fisher's *z* test for two correlations. See Table 14 for the results of the Pearson product moment correlation coefficients and the Fisher's *z* tests for two correlations for mathematics ACT scores and core GPAs disaggregated by race.

Table 14

Results for Mathematics ACT Scores and Core GPAs Disaggregated by Race

Subgroup 1	*r(df)*	Subgroup 2	*r(df)*	*z*	*p*
White	.564*** (887)	Black	.600*** (222)	-0.72	.472
White	.564*** (887)	Hispanic	.437*** (311)	2.58	.010
White	.564*** (887)	Asian	.505*** (131)	0.88	.379
White	.564*** (887)	Other	.524*** (163)	0.67	.503
Black	.600*** (222)	Hispanic	.437*** (311)	2.55	.011
Black	.600*** (222)	Asian	.505*** (131)	1.24	.215
Black	.600*** (222)	Other	.524*** (163)	1.08	.280
Hispanic	.437*** (311)	Asian	.505*** (131)	-0.84	.401
Hispanic	.437*** (311)	Other	.524*** (163)	-1.17	.242
Asian	.505*** (131)	Other	.524*** (163)	-0.22	.826

Note. $^*p < .05$, $^{**}p < .01$, $^{***}p < .001$

The correlation coefficients provided evidence for a moderate positive relationship between mathematics ACT scores and core GPAs for each race subgroup. The results of the independent-samples *t* test indicated a statistically significant relationship between mathematics ACT scores and core GPAs for each race subgroup. The results of the Fisher's *z* tests for two correlations indicated no difference between the two values when comparing white students to black students, white students to Asian students, white students to students in the other race category, black students to Asian students, black students to students in the other race category, Hispanic students to Asian students, Hispanic students to students in the other race category, and Asian students to students in the other race category. The results of the Fisher's *z* test for two correlations indicated a statistically significant difference between the two values when comparing white students to Hispanic students and when comparing black students to Hispanic students. The correlation for white students was stronger than the correlation for Hispanic students. Additionally, the correlation for black students was stronger than the correlation for Hispanic students. The correlation between English ACT scores and core GPAs was affected by race.

H13. There is a difference in the relationship between reading ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among different races.

Prior to testing H13, the sample data was disaggregated by race (white, black, Hispanic, Asian, other). Five Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between reading ACT scores and core GPAs for white, black, Hispanic, Asian, and other students (Native American, Alaskan Native, Native Hawaiian, other Pacific Islander, or Multiracial). The level of significance was set at .05. Ten Fisher's *z* tests for two correlations were conducted to determine if the differences in the relationship between reading ACT scores and core GPAs for white, black, Hispanic, Asian, and the other students were statistically significant. The level of significance was set at .05 for each Fisher's *z* test for two correlations. See Table 15 for the results of the Pearson product moment correlation coefficients and the Fisher's *z* tests for two correlations for reading ACT scores and core GPAs disaggregated by race.

Table 15

Results for Reading ACT Scores and Core GPAs Disaggregated by Race

Subgroup 1	*r(df)*	Subgroup 2	*r(df)*	*z*	*p*
White	.451*** (887)	Black	.440*** (222)	0.18	.857
White	.451*** (887)	Hispanic	.366*** (311)	1.55	.121
White	.451*** (887)	Asian	.383*** (131)	0.88	.379
White	.451*** (887)	Other	.498*** (163)	-0.71	.478
Black	.440*** (222)	Hispanic	.366*** (311)	1.00	.317
Black	.440*** (222)	Asian	.383*** (131)	0.62	.535
Black	.440*** (222)	Other	.498*** (163)	-0.72	.472
Hispanic	.366*** (311)	Asian	.383*** (131)	-0.19	.849
Hispanic	.366*** (311)	Other	.498*** (163)	-1.68	.093
Asian	.383*** (131)	Other	.498*** (163)	-1.22	.223

Note. *$p < .05$, **$p < .01$, ***$p <.001$

The correlation coefficients provided evidence for a moderate positive relationship between reading ACT scores and core GPAs for each race subgroup. The results of the independent-samples *t* test indicated a statistically significant relationship between reading ACT scores and core GPAs for each race subgroup. The results of the Fisher's *z* tests for two correlations indicated no difference when comparing any of the race subgroups. Thus, the correlation between reading ACT scores and core GPAs was not affected by race.

H14. There is a difference in the relationship between science ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among different races.

Prior to testing H14, the sample data was disaggregated by race (white, black, Hispanic, Asian, other). Five Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between science ACT scores and core GPAs for white, black, Hispanic, Asian, and other students (Native American, Alaskan Native, Native Hawaiian, other Pacific Islander, or Multiracial). The level of significance was set at .05. Ten Fisher's *z* tests for two correlations were conducted to determine if the differences in the relationship between science ACT scores and core GPAs for white, black, Hispanic, Asian, and the other students were statistically significant. The level of significance was set at .05 for each Fisher's *z* test for two correlations. See Table 16 for the results of the Pearson product moment correlation coefficients and the Fisher's *z* tests for two correlations for science ACT scores and core GPAs disaggregated by race.

Table 16

Results for Science ACT Scores and Core GPAs Disaggregated by Race

Subgroup 1	*r(df)*	Subgroup 2	*r(df)*	*z*	*p*
White	.484*** (887)	Black	.459*** (222)	0.43	.667
White	.484*** (887)	Hispanic	.460*** (311)	0.47	.638
White	.484*** (887)	Asian	.385*** (131)	1.30	.194
White	.484*** (887)	Other	.496*** (163)	-0.18	.857
Black	.459*** (222)	Hispanic	.460*** (311)	-0.01	.992
Black	.459*** (222)	Asian	.385*** (131)	0.82	.412
Black	.459*** (222)	Other	.496*** (163)	-0.46	.646
Hispanic	.460*** (311)	Asian	.385*** (131)	0.87	.384
Hispanic	.460*** (311)	Other	.496*** (163)	-0.48	.631
Asian	.385*** (131)	Other	.496*** (163)	-1.17	.242

Note. **p* < .05, ***p* < .01, ****p* <.001

The correlation coefficients provided evidence for a moderate positive relationship between science ACT scores and core GPAs for each race subgroup. The results of the independent-samples *t* test indicated a statistically significant relationship between science ACT scores and core GPAs for each race subgroup. The results of the Fisher's *z* tests for two correlations indicated no difference when comparing any of the race subgroups. Thus, the correlation between science ACT scores and core GPAs was not affected by race.

H15. There is a difference in the relationship between composite ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among different races.

Prior to testing H15, the sample data was disaggregated by race (white, black, Hispanic, Asian, other). Five Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between composite ACT scores and core GPAs for white, black, Hispanic, Asian, and other students (Native American, Alaskan Native, Native Hawaiian, other Pacific Islander, or Multiracial). The level of significance was set at .05. Ten Fisher's *z* tests for two correlations were conducted to determine if the differences in the relationship between composite ACT scores and core GPAs for white, black, Hispanic, Asian, and the other students were statistically significant. The level of significance was set at .05 for each Fisher's *z* test for two correlations. See Table 17 for the results of the Pearson product moment correlation coefficients and the Fisher's *z* tests for two correlations for composite ACT scores and core GPAs disaggregated by race.

Table 17

Results for Composite ACT Scores and Core GPAs Disaggregated by Race

Subgroup 1	*r(df)*	Subgroup 2	*r(df)*	*z*	*p*
White	.566*** (887)	Black	.556*** (222)	0.18	.857
White	.566*** (887)	Hispanic	.477*** (311)	1.55	.121
White	.566*** (887)	Asian	.467*** (131)	0.88	.379
White	.566*** (887)	Other	.572*** (163)	-0.71	.478
Black	.556*** (222)	Hispanic	.477*** (311)	1.00	.317
Black	.556*** (222)	Asian	.467*** (131)	0.62	.535
Black	.556*** (222)	Other	.572*** (163)	-0.72	.472
Hispanic	.477*** (311)	Asian	.467*** (131)	-0.19	.849
Hispanic	.477*** (311)	Other	.572*** (163)	-1.68	.093
Asian	.467*** (131)	Other	.572*** (163)	-1.22	.223

Note. *p < .05, **p < .01, ***p <.001

The correlation coefficients provided evidence for a moderate positive relationship between composite ACT scores and core GPAs for each race subgroup. The results of the independent-samples *t* test indicated a statistically significant relationship between composite ACT scores and core GPAs for each race subgroup. The results of the Fisher's *z* tests for two correlations indicated no difference when comparing any of the race subgroups. Thus, the correlation between composite ACT scores and core GPAs was not affected by race.

H16. There is a difference in the relationship between English ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on SES levels.

Prior to testing H16, the sample data was disaggregated by SES level (students who qualified for free or reduced lunch, students who qualified for full price lunch). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between English ACT scores and core GPAs. The level of significance was set at .05. The correlation coefficient (r = .406) provided evidence for a moderate positive relationship between English ACT scores and core GPAs for students who qualified for free or reduced lunch. The results of the independent-samples t test indicated a statistically significant relationship between English ACT scores and core GPAs for students who qualified for free or reduced lunch, df = 707, p = .000. The correlation coefficient (r = .561) provided evidence for a moderate positive relationship between English ACT scores and core GPAs for students who qualified for full price lunch. The results of the independent-samples t test indicated a statistically significant relationship between English ACT scores and core GPAs for students who qualified for full price lunch, df = 1013, p = .000.

A Fisher's z test for two correlations was conducted to determine if the difference in relationship between English ACT scores and core GPAs for students who qualified for free or reduced lunch and students who qualified for full price lunch was statistically significant. The sample correlation for students who qualified for free or reduced lunch was compared to the sample correlation for students who qualified for full price lunch. The level of significance was set at .05. The results of the Fisher's z test for two correlations indicated a statistically significant difference between the two values, z = -4.15, p = .000. The correlation for students who qualified for free or reduced lunch (r = .406) was weaker than the correlation for students who qualified for full price lunch

(r = .561). The correlation between English ACT scores and core GPAs was affected by socioeconomic status.

H17. There is a difference in the relationship between mathematics ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on SES levels.

Prior to testing H17, the sample data was disaggregated by SES level (students who qualified for free or reduced lunch, students who qualified for full price lunch). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between mathematics ACT scores and core GPAs. The level of significance was set at .05. The correlation coefficient (r = .479) provided evidence for a moderate positive relationship between mathematics ACT scores and core GPAs for students who qualified for free or reduced lunch. The results of the independent-samples t test indicated a statistically significant relationship between mathematics ACT scores and core GPAs for students who qualified for free or reduced lunch, df = 707, p = .000. The correlation coefficient (r = .585) provided evidence for a moderate positive relationship between mathematics ACT scores and core GPAs for students who qualified for full price lunch. The results of the independent-samples t test indicated a statistically significant relationship between mathematics ACT scores and core GPAs for students who qualified for full price lunch, df = 1013, p = .000.

A Fisher's z test for two correlations was conducted to determine if the difference in relationship between mathematics ACT scores and core GPAs for students who qualified for free or reduced lunch and students who qualified for full price lunch was statistically significant. The sample correlation for students who qualified for free or reduced lunch was compared to the sample correlation for students who qualified for full price lunch. The level of significance was set at .05. The results of the Fisher's *z* test for two correlations indicated a statistically significant difference between the two values, $z = -3.03$, $p = .002$. The correlation for students who qualified for free or reduced lunch ($r = .479$) was weaker than the correlation for students who qualified for full price lunch ($r = .585$). The correlation between mathematics ACT scores and core GPAs was affected by socioeconomic status.

H18. There is a difference in the relationship between reading ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on SES levels.

Prior to testing H18, the sample data was disaggregated by SES level (students who qualified for free or reduced lunch, students who qualified for full price lunch). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between reading ACT scores and core GPAs. The level of significance was set at .05. The correlation coefficient ($r = .336$) provided evidence for a weak positive relationship between reading ACT scores and core GPAs for students who qualified for free or reduced lunch. The results of the independent-samples *t* test indicated a statistically significant relationship between reading ACT scores and core GPAs for students who qualified for free or reduced lunch, $df = 707$, $p = .000$. The correlation coefficient ($r = .498$) provided evidence for a moderate positive relationship between reading ACT scores and core GPAs for students who qualified for full price lunch. The results of the independent-samples *t* test indicated a statistically significant relationship between reading ACT scores and core GPAs for students who qualified for full price lunch, $df = 1013$, $p = .000$.

A Fisher's z test for two correlations was conducted to determine if the difference in relationship between reading ACT scores and core GPAs for students who qualified for free or reduced lunch and students who qualified for full price lunch was statistically significant. The sample correlation for students who qualified for free or reduced lunch was compared to the sample correlation for students who qualified for full price lunch. The level of significance was set at .05. The results of the Fisher's z test for two correlations indicated a statistically significant difference between the two values, $z = -4.02$, $p = .000$. The correlation for students who qualified for free or reduced lunch ($r = .336$) was weaker than the correlation for students who qualified for full price lunch ($r = .498$). The correlation between reading ACT scores and core GPAs was affected by socioeconomic status.

H19. There is a difference in the relationship between science ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on SES levels.

Prior to testing H19, the sample data was disaggregated by SES level (students who qualified for free or reduced lunch, students who qualified for full price lunch). Two Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between science ACT scores and core GPAs. The level of significance was set at .05. The correlation coefficient ($r = .395$) provided evidence for a weak positive relationship between science ACT scores and core GPAs for students who qualified for free or reduced lunch. The results of the independent-samples *t* test indicated a statistically significant relationship between science ACT scores and core GPAs for students who qualified for free or reduced lunch, *df* = 707, *p* = .000. The correlation coefficient (*r* = .532) provided evidence for a moderate positive relationship between science ACT scores and core GPAs for students who qualified for full price lunch. The results of the independent-samples *t* test indicated a statistically significant relationship between science ACT scores and core GPAs for students who qualified for full price lunch, *df* = 1013, *p* = .000.

A Fisher's *z* test for two correlations was conducted to determine if the difference in relationship between science ACT scores and core GPAs for students who qualified for free or reduced lunch and students who qualified for full price lunch was statistically significant. The sample correlation for students who qualified for free or reduced lunch was compared to the sample correlation for students who qualified for full price lunch. The level of significance was set at .05. The results of the Fisher's *z* test for two correlations indicated a statistically significant difference between the two values, *z* = -3.57, *p* = .000. The correlation for students who qualified for free or reduced lunch (*r* = .395) was weaker than the correlation for students who qualified for full price lunch (*r* = .532). The correlation between science ACT scores and core GPAs was affected by socioeconomic status.

H20. There is a difference in the relationship between composite ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts based on SES levels.

Prior to testing H20, the sample data was disaggregated by SES level (students who qualified for free or reduced lunch, students who qualified for full price lunch). Two

Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between composite ACT scores and core GPAs. The level of significance was set at .05. The correlation coefficient (r = .450) provided evidence for a moderate positive relationship between composite ACT scores and core GPAs for students who qualified for free or reduced lunch. The results of the independent-samples t test indicated a statistically significant relationship between composite ACT scores and core GPAs for students who qualified for free or reduced lunch, df = 707, p = .000. The correlation coefficient (r = .607) provided evidence for a moderately strong positive relationship between composite ACT scores and core GPAs for students who qualified for full price lunch. The results of the independent-samples t test indicated a statistically significant relationship between composite ACT scores and core GPAs for students who qualified for full price lunch, df = 1723, p = .000.

A Fisher's z test for two correlations was conducted to determine if the difference in relationship between composite ACT scores and core GPAs for students who qualified for free or reduced lunch and students who qualified for full price lunch was statistically significant. The sample correlation for students who qualified for free or reduced lunch was compared to the sample correlation for students who qualified for full price lunch. The level of significance was set at .05. The results of the Fisher's z test for two correlations indicated a statistically significant difference between the two values, z = -2.15, p = .032. The correlation for students who qualified for free or reduced lunch (r = .450) was weaker than the correlation for students who qualified for full price lunch (r = .607). The correlation between composite ACT scores and core GPAs was affected by socioeconomic status.

H21. There is a difference in the relationship between English ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among community types.

Prior to testing H21, the sample data was disaggregated by community type (urban, suburban, and rural). Three Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between English ACT scores and core GPAs for urban, suburban, and rural community types. The level of significance was set at .05. Three Fisher's *z* tests for two correlations were conducted to determine if the differences were statistically significant. The level of significance was set at .05 for each of the Fisher's *z* tests for two correlations. See Table 18 for the results of the Pearson product moment correlation coefficients and the Fisher's *z* tests for two correlations for English ACT scores and core GPAs disaggregated by community type.

Table 18

Results for English ACT Scores and Core GPAs Disaggregated by Community Type

Subgroup 1	*r(df)*	Subgroup 2	*r(df)*	*z*	*p*
Urban	.456*** (1147)	Suburban	.658*** (478)	-5.45	.000
Urban	.456*** (1147)	Rural	.480*** (93)	-0.28	.780
Suburban	.658*** (478)	Rural	.480*** (93)	2.34	.019

Note. **p* < .05, ***p* < .01, ****p* <.001

The correlation coefficients provided evidence for a moderately strong positive relationship between English ACT scores and core GPAs for each community type subgroup. The results of the independent-samples *t* test indicated a statistically significant relationship between English ACT scores and core GPAs for each community type subgroup. The results of the Fisher's *z* tests for two correlations indicated no difference when comparing the correlation for students in an urban high school to the correlation for students in a rural high school. The results of the Fisher's *z* test for two correlations indicated a statistically significant difference between the two values when comparing the correlation for students in an urban high school to the correlation for students in a suburban high school. The correlation for students in an urban high school was weaker than the correlation for students in a suburban high school. Additionally, the results of the Fisher's *z* test for two correlations indicated a statistically significant difference between the two values when comparing the correlation for students in a suburban high school to the correlation for students in a rural high school. The correlation for students in a suburban high school was stronger than the correlation for students in a rural high school. The correlation between English ACT scores and core GPAs was affected by community type.

H22. There is a difference in the relationship between mathematics ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among community types.

Prior to testing H22, the sample data was disaggregated by community type (urban, suburban, and rural). Three Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between mathematics ACT scores and core GPAs for urban, suburban, and rural community types. The level of significance was set at .05. Three Fisher's *z* tests for two correlations were conducted to determine if the differences were statistically significant. The level of significance was set at .05 for each of the Fisher's *z* tests for two correlations. See Table

19 for the results of the Pearson product moment correlation coefficients and the Fisher's *z* tests for two correlations for mathematics ACT scores and core GPAs disaggregated by community type.

Table 19

Results for Mathematics ACT Scores and Core GPAs Disaggregated by Community Type

Subgroup 1	*r(df)*	Subgroup 2	*r(df)*	*z*	*p*
Urban	.515*** (1147)	Suburban	.644*** (478)	-5.45	.000
Urban	.515*** (1147)	Rural	.611*** (93)	-0.28	.194
Suburban	.644*** (478)	Rural	.611*** (93)	2.34	.631

Note. *$p < .05$, **$p < .01$, ***$p < .001$

The correlation coefficients provided evidence for a moderately strong positive relationship between mathematics ACT scores and core GPAs for each community type subgroup. The results of the independent-samples *t* test indicated a statistically significant relationship between mathematics ACT scores and core GPAs for each community type subgroup. The results of the Fisher's *z* tests for two correlations indicated no difference when comparing the correlation for students in an urban high school to the correlation for students in a rural high school and when comparing the correlation for students in a suburban high school to the correlation for students in a rural high school. The results of the Fisher's *z* test for two correlations indicated a statistically significant difference between the two values when comparing the correlation for students in an urban high school to the correlation for students in a suburban high school. The correlation for students in an urban high school was weaker than the correlation for students in a suburban high school. The correlation between mathematics ACT scores and core GPAs was affected by community type.

H23. There is a difference in the relationship between reading ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among community types.

Prior to testing H23, the sample data was disaggregated by community type (urban, suburban, and rural). Three Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between reading ACT scores and core GPAs for urban, suburban, and rural community types. The level of significance was set at .05. Three Fisher's *z* tests for two correlations were conducted to determine if the differences were statistically significant. The level of significance was set at .05 for each of the Fisher's *z* tests for two correlations. See Table 20 for the results of the Pearson product moment correlation coefficients and the Fisher's *z* tests for two correlations for reading ACT scores and core GPAs disaggregated by community type.

Table 20

Results for Reading ACT Scores and Core GPAs Disaggregated by Community Type

Subgroup 1	*r(df)*	Subgroup 2	*r(df)*	*z*	*p*
Urban	.385*** (1147)	Suburban	.604*** (478)	-5.39	.000
Urban	.385*** (1147)	Rural	.414*** (93)	-0.32	.749
Suburban	.604*** (478)	Rural	.414*** (93)	2.27	.023

Note. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

The correlation coefficients provided evidence for a weak positive relationship between reading ACT scores and core GPAs for students in an urban high school, moderately strong positive relationship for students in a suburban high school, and moderate positive relationship for students in a rural high school. The results of the independent-samples *t* test indicated a statistically significant relationship between reading ACT scores and core GPAs for each community type subgroup. The results of the Fisher's *z* tests for two correlations indicated no difference when comparing the correlation for students in an urban high school to the correlation for students in a rural high school. The results of the Fisher's *z* tests for two correlations indicated no difference when comparing the correlation for students in an urban high school to the correlation for students in a rural high school. The results of the Fisher's *z* test for two correlations indicated a statistically significant difference between the two values when comparing the correlation for students in an urban high school to the correlation for students in a suburban high school. The correlation for students in an urban high school was weaker than the correlation for students in a suburban high school. Additionally, the results of the Fisher's *z* test for two correlations indicated a statistically significant difference between the two values when comparing the correlation for students in a suburban high school to the correlation for students in a rural high school. The correlation for students in a suburban high school was stronger than the correlation for students in a rural high school. The correlation between reading ACT scores and core GPAs was affected by community type.

H24. There is a difference in the relationship between science ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among community types.

Prior to testing H24, the sample data was disaggregated by community type (urban, suburban, and rural). Three Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between science ACT scores and core GPAs for urban, suburban, and rural community types. The level of significance was set at .05. Three Fisher's *z* tests for two correlations were conducted to determine if the differences were statistically significant. The level of significance was set at .05 for each of the Fisher's *z* tests for two correlations. See Table 21 for the results of the Pearson product moment correlation coefficients and the Fisher's *z* tests for two correlations for science ACT scores and core GPAs disaggregated by community type.

Table 21

Results for Science ACT Scores and Core GPAs Disaggregated by Community Type

Subgroup 1	*r(df)*	Subgroup 2	*r(df)*	*z*	*p*
Urban	.431*** (1147)	Suburban	.623*** (478)	-4.93	.000
Urban	.431*** (1147)	Rural	.497*** (93)	-0.78	.435
Suburban	.623*** (478)	Rural	.497*** (93)	1.62	.105

Note. **p* < .05, ***p* < .01, ****p* <.001

The correlation coefficients provided evidence for a moderately strong positive relationship between science ACT scores and core GPAs for each community type subgroup. The results of the independent-samples *t* test indicated a statistically significant relationship between science ACT scores and core GPAs for each community type subgroup. The results of the Fisher's *z* tests for two correlations indicated no difference when comparing the correlation for students in an urban high school to the correlation for students in a rural high school and when comparing the correlation for students in a suburban high school to the correlation for students in a rural high school. The results of the Fisher's *z* test for two correlations indicated a statistically significant difference between the two values when comparing the correlation for students in an urban high school to the correlation for students in a suburban high school. The correlation for students in an urban high school was weaker than the correlation for students in a suburban high school. The correlation between science ACT scores and core GPAs was affected by community type.

H25. There is a difference in the relationship between composite ACT scores and core GPAs for students in grades 9 through 12 in three south-central Kansas districts among community types.

Prior to testing H25, the sample data was disaggregated by community type (urban, suburban, and rural). Three Pearson product moment correlation coefficients were calculated to determine the strength and direction of the relationship between composite ACT scores and core GPAs for urban, suburban, and rural community types. The level of significance was set at .05. Three Fisher's *z* tests for two correlations were conducted to determine if the differences were statistically significant. The level of significance was set at .05 for each of the Fisher's *z* tests for two correlations. See Table 22 for the results of the Pearson product moment correlation coefficients and the Fisher's *z* tests for two correlations for composite ACT scores and core GPAs disaggregated by community type.

Table 22

Results for Composite ACT Scores and Core GPAs Disaggregated by Community Type

Subgroup 1	*r(df)*	Subgroup 2	*r(df)*	*z*	*p*
Urban	.495*** (1147)	Suburban	.707*** (478)	-6.21	.000
Urban	.495*** (1147)	Rural	.561*** (93)	-0.85	.395
Suburban	.707*** (478)	Rural	.561*** (93)	2.17	.030

Note. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

The correlation coefficients provided evidence for a moderate positive relationship between composite ACT scores and core GPAs for students in an urban high school, strong positive relationship for students in a suburban high school, and moderately strong positive relationship for students in a rural high school. The results of the independent-samples *t* test indicated a statistically significant relationship between composite ACT scores and core GPAs for each community type subgroup. The results of the Fisher's *z* tests for two correlations indicated no difference when comparing the correlation for students in an urban high school to the correlation for students in a rural high school. The results of the Fisher's *z* tests for two correlations indicated no difference when comparing the correlation for students in an urban high school to the correlation for students in a rural high school. Thus, the correlation between composite ACT scores and core GPAs was not affected by community type when comparing students in an urban high school to students in a rural high school. The results of the Fisher's *z* test for two correlations indicated a statistically significant difference between the two values when comparing the correlation for students in an urban high school to the correlation for students in a suburban high school. The correlation for students in an urban high school was weaker than the correlation for students in a suburban high school. Additionally, the results of the Fisher's *z* test for two correlations indicated a statistically significant difference between the two values when comparing the correlation for students in a suburban high school to the correlation for students in a rural high school. The correlation for students in a suburban high school was stronger than the correlation for students in a rural high school. The correlation between composite ACT scores and core GPAs was affected by community type.

Summary

The results from research question one and two indicate that ACT scores and core GPAs have a moderate positive relationship. Reading ACT scores have the weakest correlation to core GPAs, while composite ACT scores have the strongest correlation to core GPAs. Gender did not affect the correlation between ACT scores (English, mathematics, reading, science, composite) and core GPAs. Race only affected the correlation between ACT scores and core GPAs when comparing the correlations for white students and Hispanic students on the English ACT and mathematics ACT, and when comparing correlations between black students and Hispanic students on the mathematics ACT. The results also indicated that socioeconomic status affects the correlation between ACT scores (English, mathematics, reading, science, composite) and core GPAs. The correlation between ACT scores (English, mathematics, reading, science, composite) and core GPAs is stronger for students who qualified for full price lunch than for students who qualified for free or reduced lunch. Community type also affected the correlation between ACT scores and core GPAs. The correlation between ACT scores (English, mathematics, reading, science, composite) and core GPAs for students in an urban high school were weaker than the correlation for students in a suburban high school. Students in rural high schools had weaker correlations between ACT scores (English, reading, and composite) and core GPAs than students in suburban high schools.

Chapter 4 focused on the results of the study, including descriptive statistics and the outcomes of the hypothesis tests. The next chapter presents a study summary, which includes an overview of the problem, purpose statement, research questions, review of the methodology, and major findings. In addition, findings related to the literature, implications for action, recommendations for future research, and concluding remarks are included in Chapter 5.

Chapter 5

Interpretation and Recommendations

With education hitting the national spotlight, there has been increasing pressure for educators to define the validity of grades students earn and to increase performance on national standardized tests (Zhang & Sanchez, 2013). The ACT, composed of four tests (English, mathematics, reading, and science), was designed to measure a student's capacity to succeed in rigorous post-secondary courses and to determine the academic skills a student developed via secondary coursework (ACT, 2005). This study was conducted to determine the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs for high school graduates in the class of 2017 in three south-central Kansas districts. Furthermore, the study was conducted to determine whether the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs was affected by gender, race, socioeconomic status, and high school community type. Chapter 5 includes a summary of the study, including an overview of the problem, purpose statement and research questions, review of the methodology and major findings. Also, findings related to the literature, implications for action, recommendations for future research, and concluding remarks are all incorporated in this chapter.

Study Summary

The average ACT score in the state of Kansas is 21.7, which indicates the typical Kansas high school graduate is college ready (KSDE, 2017a, 2017b, 2017c). However, there is a large discrepancy in districts across Kansas (KSDE, 2017a, 2017b, 2017c). Despite an average ACT score which indicates the typical Kansas high school graduate is college ready, in 2015, McDaniel noted that over 27% of high school graduates were required to enroll in remedial post-secondary coursework. This study reported the correlation between ACT scores (English, mathematics, reading, science, composite) and core GPAs. Whether the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs was affected by students' gender, race, socioeconomic status, and community type was also studied. Provided in the following sections are an overview of the problem, purpose statement and research questions, review of the methodology, and major findings.

Overview of the problem. The landscape of American education has changed drastically since the 19th century, but the original grading paradigm is still prevalent in today's schools (Vatterott, 2015). Since ACT scores and GPAs are significant factors in student post-secondary acceptance and retention, increasing attention has been given to the validity of grades (Pope, 2011). Researchers have increasingly used the correlation between ACT scores and GPAs as their basis for claims for or against the validity of grades, without first establishing the original nature of the relationship between ACT scores and GPAs (Woodruff & Ziomek, 2004; Zhang & Sanchez, 2013). Therefore, it was important to determine a baseline correlation for the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs. Additionally, there was a need to determine whether students' gender, race, socioeconomic status, or community type affected the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs.

Purpose statement and research questions. This study was conducted to examine the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs calculated from all English, mathematics, history, and science courses in grades 9 through 12 in high schools in three south-central Kansas school districts. An additional purpose of the study was to explore how gender, race, socioeconomic status, and high school community type affect the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs for students in grades 9 through 12 in three south-central Kansas districts. Two research questions with 25 hypotheses were developed and tested to address the purposes of this study.

Review of the methodology. A quantitative correlational design was used in this study, which utilized archived student data from students who took the ACT and graduated from the Sunflower, Marigold, and Daisy School Districts in 2017. The variables of interest in this study were ACT score and core GPA. Additionally, gender, race, socioeconomic status, and community type were examined to determine if the demographics affected the relationship between ACT score and core GPA. Sixty-five Pearson product correlations, 65 independent-samples *t* tests, and 75 Fisher's *z* tests for two correlations were conducted to respond to the two research questions and 25 hypotheses.

Major findings. The results of the data analysis indicated that students' ACT scores and core GPAs have a moderately strong correlation. Student gender did not affect the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs. The relationship between English ACT scores and core GPAs was significantly stronger for white students than Hispanic students. Additionally, the relationship between mathematics ACT scores and core GPAs was significantly stronger for white and black students than Hispanic students. Student race did not affect the relationship between ACT scores and core GPAs for reading, science, or composite scores. The relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs was significantly weaker for students who qualified for free or reduced lunch than for students who qualified for full price lunch. Furthermore, students in an urban high school had a significantly weaker relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs than students in a suburban high school. Lastly, the results indicated the relationship between ACT scores and core GPAs was significantly stronger for students in a suburban high school than students in a rural high school for English, reading, and composite scores.

Findings Related to the Literature

This section assesses the current study's findings as they relate to the literature regarding the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs and whether the relationship is affected by gender, race, socioeconomic status, and high school community type. According to Brookhart (2016), over 100 years of research has consistently determined the correlation coefficient for the relationship between standardized test scores and core GPAs has typically been approximately $r = .50$. The results of this study demonstrated a moderately strong positive relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs. The correlations ranged from $r = .478$ to $r = .584$, which is consistent with the studies in the literature review (Brennan et al., 2001; Carter, 1952; Duckworth et al., 2012; Duckworth & Seligman, 2006; McCandless et al., 1972; Moore, 1939; Pattison et al., 2013; Ross & Hooks, 1930; Unzicker, 1925; Woodruff & Ziomek, 2004).

Ross and Hooks (1929), McCandless et al. (1972), Unzicker (1925), Carter (1952), Brennan et al. (2001), Woodruff and Ziomek (2004), and Pattison et al. (2013) all found moderate positive correlations in their studies. Ross and Hooks (1929) found a moderate positive correlation between intelligence test scores and GPAs for elementary, middle, and high school students when completing a meta-analysis. McCandless et al.'s (1972) study of the relationship between Algebra achievement scores and GPAs in 433 students also yielded a moderate positive correlation. Unzicker (1925) found a moderate positive correlation coefficient during his study on the relationship between standardized achievement tests and GPAs. Carter (1952) correlated Algebra achievement scores and GPAs of 235 high school students with a moderate positive correlation. Brennan et al. (2001) compared the Massachusetts standardized state reading test scores to mathematics, English, and science grades of 736 eighth-grade students and found moderate positive correlations. Woodruff and Ziomek (2004) compared composite ACT scores and GPA for every high school student who took the ACT between 1991 and 2003. The analysis, consisting of a sample size of nearly 700,000 unique students, found moderate positive correlations for self-reported average GPAs and composite ACT scores, self-reported mathematics grades and ACT mathematics scores, and self-reported English grades and ACT English scores (Woodruff & Ziomek, 2004). Pattison et al. (2013) compiled a sample size of over 40,000 students' GPA and ACT content scores spanning 1972 to 2002, with moderately strong correlations for both reading and mathematics ACT sub-scores (Pattison et al., 2013). The results from this study are consistent with the findings from Ross and Hooks (1929), McCandless et al. (1972), Unzicker (1925), Carter (1952), Brennan et al. (2001), Woodruff and Ziomek (2004), and Pattison et al. (2013).

Moore (1939), Duckworth and Seligman (2006), and Duckworth et al. (2012) all found moderately strong positive correlations in their studies. Moore (1939) conducted a correlation study between student GPAs and Stanford Intelligence Test scores with a moderately strong positive correlation. Duckworth and Seligman (2006) found a moderately strong positive correlation between eighth-grade student GPAs and TerraNova California assessment scores in a correlation study. Similarly, Duckworth et al. (2012) found a moderately strong positive correlation between reading and mathematics ACT scores and student GPAs. Similar to the current study, all three results indicate moderately strong positive correlations between standardized test scores and core GPAs regardless of the type of standardized test.

Over the last two decades, several studies have been conducted regarding the variables that affect the relationship between ACT scores and core GPAs (Bolek, 2011; Ginstead, 2013; Roth et al., 2000; Schiel et al., 1996). The findings from the current study refute the research from Schiel et al. (1996), Roth et al. (2000), Bolek (2011), and Ginstead (2013) described in the literature review. Schiel et al. (1996) and Roth et al. (2000) both reported that the relationship between student grades and ACT Plan scores were not affected by race or socioeconomic status. Bolek (2011) also suggested that socioeconomic status did not affect the relationship between ACT scores and student GPAs. In the current study, the correlation between ACT scores and core GPAs in all four subject area tests and the composite ACT score for students who qualified for free or reduced lunch was significantly lower than the correlation for students who qualified for full price lunch. Additionally, the relationship between English ACT scores and core GPAs was significantly stronger for white students than Hispanic students. Furthermore, the relationship between mathematics ACT scores and core GPAs for white and black students was significantly stronger than for Hispanic students in the present study. The results of the current study indicated gender did not affect the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs. This contrasts with Ginstead (2013) who reported a significantly weaker correlation between mathematics ACT scores and GPAs for females than for males.

Conclusions

As community pressure for high graduation rates and positive ACT results mounts, school districts have funneled money into preparing students for the test and raising students' academic profiles. The results of the current study indicated that ACT scores (English, mathematics, reading, science, composite) and core GPAs are moderately correlated, and the student demographics of race, socioeconomic status, and community type affect the strength of the correlation. Building leaders should respond to the results of this study by focusing on professional development that allows teachers to provide all students with ample college or career preparation. Future research should be expanded to more districts including districts with an SBG paradigm, examine teachers' criteria in determining grades, factors that confound student ACT scores and GPAs, and how to use ACT scores and GPAs to predict college success. School districts should work together to provide access to ACT preparation materials and meaningful grading practices.

Implications for action. The data indicated that ACT scores (English, mathematics, reading, science, composite) and core GPAs have a moderate correlation. Grades are multi-dimensional measures of student performance, whereas standardized tests are meant to measure student skill in specific content areas. The findings from this study have implications for district-level and building-level administrators to foster a review of grading practices and promote conversations about the validity of using standardized test scores and GPAs to measure grading practices.

The current study provided evidence that there could be more grade inflation for Hispanic students since the relationship between English and mathematics ACT scores and core GPAs is significantly weaker for Hispanic students than for white or black students. Furthermore, the data from the study indicated that students who qualified for free or reduced lunch had a significantly weaker relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs than students who qualified for full price lunch. The findings from this study should encourage district and building leaders and teachers to examine the discrepancy in correlation between ACT scores and core GPAs based on socioeconomic status and community type and expand professional development models related to grading practices and ACT preparation.

The data indicated that students who attend suburban high schools have core GPAs that have a significantly stronger relationship with ACT scores, thus potentially less grade inflation, than students who attend urban or rural high schools. The results of the current study have implications for building and district leaders to work with teachers inside and outside the district to understand how changes in instructional design lead to changes in grading practices, grade distributions, and the usefulness of grades as predictors for future educational outcomes. The Sunflower and Marigold School Districts should provide professional development and training for all staff on strategies to enhance the rigor of all classes, so all students are exposed to curricula that provides ample preparation for the college or career pathway regardless of the geographic location of the high school. Professional development for building and district administrators should include how to build productive relationships with students, families, and the community to create a culture of learning that extends past the school day bell.

Recommendations for future research. Future research in this area should be expanded to more than three districts to determine if the trends exhibited in this study also exist on a larger scale. An expanded study would generate a larger sample size, particularly for students in a rural high school, and give more credibility to the correlation between ACT scores (English, mathematics, reading, science, composite) and core GPAs. The study could also be extended to multiple graduating classes to examine data over time.

The second recommendation is to conduct a similar study involving students' ACT scores and core GPAs from districts that utilize standards-based grading (SBG). Presently, there are no districts utilizing SBG in south-central Kansas. Comparing the correlation between ACT scores (English, mathematics, reading, science, composite) and core GPAs from various grading paradigms (traditional, plus minus traditional, SBG) would allow researchers to determine if a certain grading paradigm had a strong relationship to standardized test scores.

Additionally, future research should examine the criteria teachers use in grading, teacher skill at identifying student level based on written state standards, and teacher ability to merge instructional and assessment skills (Brookhart, 2016). This investigation would allow researchers to focus on the sources of grading decisions and determine if teachers feel pressure to pass students who do not reach expected levels of achievement. Research in this area will contribute to enhancing the validity and fairness of grading practices (Brookhart, 2016).

The fourth recommendation is to conduct a qualitative study to survey students on the factors that contribute to their core GPAs and affect their ACT scores. The study would allow researchers to interview students, parents, teachers, and staff to gain valuable insight on the barriers that help or hinder student success in school or on standardized tests. Such a study would benefit school districts by developing responses to the barriers to success as perceived by students, parents, teachers, and staff.

Furthermore, future research could also extend the study to success in college. Determining whether ACT scores and core GPAs are predictors of college success, and how that relationship is affected by gender, race, socioeconomic status, and community type will allow secondary and post-secondary institutions to work in concert to align systems and work together to ensure students have the support needed for a seamless transition. This research would assist secondary schools in preparing students for college and provide post-secondary schools a better understanding of the needs of incoming students.

Concluding remarks. While this study did not support the stereotype that gender impacts the relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs, the results indicated that race, socioeconomic status, and high school community type all affect the relationship significantly. Due to the significant differences in correlation strength between subgroups, Sunflower, Daisy, and Marigold Public Schools have an obligation to respond to the results of this study. Specific and intentional district-wide beliefs about the ability to help all students achieve success in the classroom and on the ACT is the first step toward creating a stronger relationship between ACT scores (English, mathematics, reading, science, composite) and core GPAs.

References

Abbott, W. M. (2008). The politics of grade inflation: A case study. *Change: The Magazine of Higher Learning, 40*(1), 32-37. Retrieved from http://web.b.ebscohost.com/ehost/detail/detail?vid=3&sid=cb1f71f9-7ac6-47c1-9820-4ad53a0288ed%40sessionmgr120&hid=125&bdata=JnNpdGU9ZWhvc3QtbGl2ZQ%3d%3d#AN=EJ782170&db=eric

ACT. (2005). Are high school grades inflated? *Issues in College Readiness, 1*(1), 1-4. Retrieved from https://forms.act.org/research/policymakers/pdf/issues.pdf

ACT. (2014). *Technical manual*. Retrieved from https://www.act.org/content/dam/act/unsecured/documents/ACT_Technical_Manual.pdf

ACT. (2015). *ACT research & policy: Influence of achievement in core high school courses on ACT scores*. Retrieved from http://www.act.org/content/dam/act/unsecured/documents/2015-Tech-Brief-Influence-of-Achievement.pdf

ACT. (2016). *ACT history*. Retrieved from http://www.act.org/content/act/en/about-act/history.html.

ACT. (2017a). *ACT annual report 2017*. Retrieved from http://www.act.org/content/dam/act/unsecured/documents/AnnualReport2017.pdf.

ACT. (2017b). *The condition of college and career readiness 2017 – profile report - Kansas*. Retrieved from http://www.act.org/content/dam/act/unsecured/documents/cccr2017/P_17_179999_S_S_N00_ACT-GCPR_Kansas.pdf

Alexander, K. L., Entwisle, D. R., & Kabbani, N. S. (2001). The dropout process in life course perspective: Early risk factors at home and school. *The Teachers College Record, 103*(1), 760–822. doi:10.1111/0161- 4681.00134

Allen, J. D. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 78*(5), 218-223. doi:10.3200/TCHS.78.5.218-223

Allensworth, E. M. (2013). The use of ninth-grade early warning indicators to improve Chicago schools. *Journal of Education for Students Placed at Risk (JESPAR), 18*(1), 68–83. doi:10.1080/10824669.2013.745181

Allensworth, E. M., & Easton, J. Q. (2007). *What matters for staying on-track and graduating in Chicago public high schools: A close look at course grades, failures, and attendance in the freshman year*. Chicago, IL: University of Chicago Consortium on Chicago School Research.

Allensworth, E. M., Gwynne, J. A., Moore, P., & de la Torre, M. (2014). *Looking forward to high school and college: Middle grade indicators of readiness in Chicago Public Schools*. Chicago, IL: University of Chicago Consortium on Chicago School Research.

Aronson, M. J. (2008). *How teachers' perceptions in the areas of student behavior, attendance and student personality influence their grading practice* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global: The Humanities and Social Sciences Collection. (ProQuest No. 304510267)

Ashbaugh, E. J. (1924). Reducing the variability of teachers’ marks. *Journal of Educational Research, 9*(3), 185–198. doi:10.1080/00220671.1924.10879447

Atkinson, R. C., & Geiser, S. (2009). Reflections on a century of college admissions tests. *Educational Researcher, 38*(9), 665–676. doi:10.3102/0013189x09351981

Balfanz, R., Herzog, L., & MacIver, D. J. (2007). Preventing student disengagement and keeping students on the graduation path in urban middle-grades schools: Early identification and effective interventions. *Educational Psychologist, 42*(4), 223–235. doi:10.1080/00461520701621079

Banker, H. J. (1927). The significance of teachers' marks. *The Journal of Educational Research, 16*(5), 159–171. doi:10.1080/00220671.1927.10879778

Barrington, B. L., & Hendricks, B. (1989). Differentiating characteristics of high school graduates, dropouts, and non-graduates. *Journal of Educational Research, 82*(6), 309–319. doi:10.1080/00220671.1989.10885913

Beatty, A. S., Walmsley, P. T., Sackett, P. R., Kuncel, N. R., & Koch, A. J. (2015). The reliability of college grades. *Educational Measurement: Issues and Practice, 34*(4), 31–40. doi:10.1111/emip.12096

Beecher, M., & Fisher, L. (1999, Spring/Summer). High school courses and scores as predictors of college success. *Journal of College Admissions, 163*(28), 4-9. Retrieved from http://eric.ed.gov/?id=EJ592749

Bisesi, T., Farr, R., Greene, B., & Haydel, E. (2000). Reporting to parents and the community. In E. Trumbull & B. Farr (Eds.), *Grading and reporting student progress in an age of standards* (pp. 157–184). Norwood, MA: Christopher-Gordon.

Bloom, B. S. (1971). Mastery learning. In J. H. Block (Ed.), *Mastery learning: Theory and practice* (pp. 47–63). New York, NY: Holt, Rinehart & Winston.

Bolek, C. Y. (2011). *A study of discrepancies between high school students' grades and standardized test scores* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global: The Humanities and Social Sciences Collection. (UMI No. 3489812)

Bolton, F. E. (1927). Do teachers' marks vary as much as supposed? *Education, 48*(2), 28–39. Retrieved from https://files.eric.ed.gov/fulltext/ED540189.pdf

Bonner, S. M. (2016). Teacher perceptions about assessment: Relationship to practice and policy. In G. T. Brown, & L. Harris (Eds.), *Human factors and social conditions of assessment.* London, United Kingdom: Routledge.

Bonner, S. M., & Chen, P. P. (2009). Teacher candidates' perceptions about grading and constructivist teaching. *Educational Assessment*, *14*(2), 57–77. doi:10.1080/10627190903039411

Bowers, A. J. (2009). Reconsidering grades as data for decision making: More than just academic knowledge. *Journal of Educational Administration, 47*(5), 609–629. doi:10.1108/09578230910981080

Bowers, A. J. (2010). Grades and graduation: A longitudinal risk perspective to identify student dropouts. *Journal of Educational Research, 103*(3), 191– 207. doi:10.1080/00220670903382970

Bowers, A. J. (2011). Student risk factors. In D. J. Brewer & L. O. Picus (Eds.), *Encyclopedia of education economics and finance*. Thousand Oaks, CA: Sage.

Bowers, A. J. (2014). Student risk factors. In D. J. Brewer & L. O. Picus (Eds.), *Encyclopedia of education economics and finance.* Thousand Oaks, CA: Sage.

Bowers, A. J., & Sprott, R. (2012). Examining the multiple trajectories associated with dropping out of high school: A growth mixture model analysis. *Journal of Educational Research, 105*(3), 176–195. doi:10.1080/00220671.2011.552075

Bowers, A. J., Sprott, R., & Taff, S. (2013). Do we know who will drop out? A review of the predictors of dropping out of high school: Precision, sensitivity and specificity. *The High School Journal, 96*(2), 77–100. doi:10.1353/hsj.2013.0000

Brennan, R. T., Kim, J., Wenz-Gross, M., & Siperstein, G. N. (2001). The relative equitability of high-stakes testing versus teacher-assigned grades: An analysis of the Massachusetts Comprehensive Assessment System (MCAS). *Harvard Educational Review, 71*(2), 173– 215. doi:10.17763/haer.71.2.v51n6503372t4578

Brimi, H. M. (2011). Reliability of grading high school work in English. *Practical Assessment, Research & Evaluation, 16*(17). Retrieved from http://pareonline. net/getvn.asp?v=16&n=17

Brookhart, S. M. (1993). Teachers' grading practices: Meaning and values. *Journal of Educational Measurement, 30*(2), 123–142. doi:10.1111/j.17453984.1993.tb01070.x

Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education, 7*(4), 279– 301. doi:10.1207/s15324818ame0704_2 B

Brookhart, S. M. (2009). *Grading* (2nd ed.). New York, NY: Merrill Pearson Education

Brookhart, S. M. (2011). *Grading and learning: Practices that support student achievement.* Bloomington, IN: Solution Tree Press.

Brookhart, S. M. (2015). Graded achievement, tested achievement, and validity. *Educational Assessment, 20*(4), 268–296. doi:10.1080/10627197.2015.1093928

Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., …Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research, 86*(4), 803-848. doi:10.3102/0034654316672069

Cairns, R. B., Cairns, B. D., & Neckerman, H. J. (1989). Early school dropout: Configurations and determinants. *Child Development, 60*(6), 1437–1452. doi:10.2307/1130933

Canady, R. L., & Hotchkiss, P. R. (1989). It's a good score! Just a bad grade. *Phi Delta Kappan, 1*(1), 68-71. Retrieved from https://eric.ed.gov/?id=EJ395078

Carroll, J. (1963). A model of school learning. *The Teachers College Record*, *64*(8), 723–723. Retrieved from http://garfield.library.upenn.edu/classics1982/A1982NF37300001.pdf

Carter, R. S. (1952). How invalid are marks assigned by teachers? *Journal of Educational Psychology, 43*(4), 218–228. doi:10.1037/h0061688

Caruth, D. L., & Caruth, G. D. (2013). Grade inflation: An issue for higher education. *Turkish Online Journal of Distance Education, 14*(1), 102-110. doi:10.1111/emip.12107

Casillas, A., Robbins, S., Allen, J., Kuo, Y. L., Hanson, M. A., & Schmeiser, C. (2012). Predicting early academic failure in high school from prior academic achievement, psychosocial characteristics, and behavior. *Journal of Educational Psychology, 104*(2), 407–420. doi:10.1037/a0027180

Cizek, G. L. (2000). Pockets of resistance in the assessment revolution. *Educational Measurement: Issues and Practice, 19*(2), 16–23. doi:10.1111/j.1745-3992.2000.tb00026.x

Cizek, G. J., Fitzgerald, J. M., & Rachor, R. A. (1995). Teachers' assessment practices: Preparation, isolation, and the kitchen sink. *Educational Assessment, 3*(2), 159–179. doi:10.1207/s15326977ea0302_3

Clarridge, P. B. & Whitaker, E. M. (1994). Implementing a new elementary progress report. Educational Leadership, *52*(2), 7–9. Retrieved from http://www.ascd.org/publications/educationalleadership/oct94/vol52/num02/Implementing-a-NewElementary-Progress-Report.aspx

Cliffordson, C. (2008). Differential prediction of study success across academic programs in the Swedish context: The validity of grades and tests as selection instruments for higher education. *Educational Assessment, 13*(1), 56–75. doi:10.1080/10627190801968240

Cox, K. B. (2011). Putting classroom grading on the table, a reform in progress. *American Secondary Education, 40*(1), 67–87. Retrieved from https://eric.ed.gov/?id=EJ951229

Coyle, T. R., & Pillow, D. R. (2008). SAT and ACT predict college GPA. *Intelligence 36*(6), 719-729. doi:10.1016/j.intell.2008.05.001

Creason, N. (2013, September 19). New grading system for schools a step in the right direction. *The Sentinel, 1*(3), E2. Retrieved from http://thesentinel.com/

Crooks, A. D. (1933). Marks and marking systems: A digest. *Journal of Educational Research, 27*(4), 259–272. doi:10.1080/00220671.1933.10880402

Cross, L. H., & Frary, R. B. (1999). Hodgepodge grading: Endorsed by students and teachers alike. *Applied Measurement in Education*, *12*(1), 53–72. doi:10.1207/s15324818ame1201_4

Cunha, F., & Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and non-cognitive skill formation. *Journal of Human Resources, 43*(4), 738–782. doi:10.3368/jhr.43.4.738

Curreton, L. W. (1971). The history of grading practices. *National Council on Measurement in Education*, *2*(4), 1-9. Retrieved from http://tenure.umatter2.us/wp-content/uploads/2011/09/1971-Curreton-History-of-Grading.pdf

Daisy Public Schools. (2017). *2017 district profile*. Retrieved from http://www.usd[redacted].com/education/components/scrapbook/default.php?sectiondetailid=13362&

Duckworth, A. L., Quinn, P. D., & Tsukayama, E. (2012). What No Child Left Behind leaves behind: The roles of IQ and self-control in predicting standardized achievement test scores and report card grades. *Journal of Educational Psychology*, *104*(2), 439-451. doi: 10.1037/a0026280

Duckworth, A. L., & Seligman, M. E. P. (2006). Self-discipline gives girls the edge: Gender in self-discipline, grades, and achievement test scores. *Journal of Educational Psychology, 98*(1), 198–208. doi:10.1037/0022-0663.98.1.198

Duncan, R. C., & Noonan, B. (2007). Factors affecting teachers' grading and assessment practices. *Alberta Journal of Educational Research*, *53*(1), 1–21. Retrieved from https://eric.ed.gov/?id=EJ762241

Durm, M. W. (1993, Spring). An A is not an A: A history of grading. *The Educational Forum*, *57*(3), 1-4. doi:10.1080/00131729309335429

Edgeworth, F. Y. (1888). The statistics of examinations. *Journal of the Royal Statistical Society, 51*(1), 599-635. Retrieved from http://journals.sagepub.com/doi/abs/10. 1177/001316446002000110

Edwards, S. (2009). *Grade inflation of college bound students in southwest Missouri high schools* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global: The Humanities and Social Sciences Collection. (UMI No. 3361694)

Eells, W. C. (1930). Reliability of repeated grading of essay type examinations. *Journal of Educational Psychology*, *21*(1), 48-52. doi:10.1037/h0071103

Ekstrom, R. B., Goertz, M. E., Pollack, J. M., & Rock, D. A. (1986). Who drops out of high school and why? Findings from a national study. *Teachers College Record*, *87*(3), 356–373. Retrieved from https://eric.ed.gov/?id=EJ332413

Ensminger, M. E., & Slusarcick, A. L. (1992). Paths to high school graduation or dropout: A longitudinal study of a first-grade cohort. *Sociology of Education, 65*(2), 91–113. doi:10.2307/2112677

Evans, F. B. (1976). What research says about grading. In S. B. Simon & J. A. Bellanca (Eds.), *Degrading the grading myths: A primer of alternatives to grades and marks* (pp. 30–50). Washington, DC: Association for Supervision and Curriculum Development.

Farr, B. P. (2000). Grading practices: An overview of the issues. In E. Trumbull & B. Farr (Eds.), *Grading and reporting student progress in an age of standards* (pp. 1-22). Norwood, MA: Christopher-Gordon.

Farkas, G., Grobe, R. P., Sheehan, D., & Shaun, Y. (1990). Cultural resources and school success: Gender, race, and poverty groups within an urban school district. *American Sociological Review, 55*(1), 127–142. doi:10.2307/2095708

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research, 59*(2), 117–142. doi:10.3102/00346543059002117

Fitzsimmons, S. J., Cheever, J., Leonard, E., & Macunovich, D. (1969). School failures: Now and tomorrow. *Developmental Psychology, 1*(2), 134–146. doi:10.1037/h0027088

Folzer-Napier, S. (1976). Grading and young children. In S. B. Simon & J. A. Bellanca (Eds.), *Degrading the grading myths: A primer of alternatives to grades and marks* (pp. 23–27). Washington, DC: Association for Supervision and Curriculum Development.

Galton, D. J., & Galton, C. J. (1998). Francis Galton: and eugenics today. *Journal of Medical Ethics, 24*(2), 99–105. doi:10.1136/jme.24.2.99

Garton, B. L., Ball, A. L., & Dyer, J. E. (2002). The academic performance and retention of college agriculture students. *Journal of Agricultural Education, 43*(1), 46-56. Retrieved from http://pubs.aged.tamu.edu/jae/pdf/vol43/43-01-46.pdf

Geiser, S., & Santelices, M. V. (2007). Validity of high-school grades in predicting student success beyond the freshman year: High-school record vs. standardized tests as indicators of four-year college outcomes. *Center for Studies in Higher Education Research and occasional paper series: CSHE.6-07*. Berkeley, CA: University of California, Berkeley.

Geverdt, D. (2015). *Education Demographic and Geographic Estimates program (EDGE): Locale boundaries user's manual (NCES 2016-012)*. Retrieved from https://nces.ed.gov/programs/edge/docs/NCES_LOCALE_USERSMANUAL_2016012.pdf

Ginstead, M. L. (2013). Which advanced mathematics courses influence ACT scores? A state level analysis of the Iowa class of 2012. Retrieved from ProQuest Dissertations and Theses Global: The Humanities and Social Sciences Collection. (UMI No. 13622).

Glaser, R. (1963). Instructional technology and the measurement of learning outcomes: Some questions. *American Psychologist, 18*(8), 519. doi:10.1111/j.1745-3992.1994.tb00561.x

Gleason, P., & Dynarski, M. (2002). Do we know whom to serve? Issues in using risk factors to identify dropouts. *Journal of Education for Students Placed at Risk, 7*(1), 25–41. doi:10.1207/S15327671ESPR0701_3

Goodwin, B. (2011). Grade inflation: Killing with kindness? *Educational Leadership, 69*(3), 80-81. Retrieved from http://www.ascd.org/publications/educational-leadership/nov11/vol69/num03/Grade-Inflation@-Killing-with-Kindness¢.aspx

Grimes, T. V. (2010). *Interpreting the meaning of grades: A descriptive analysis of middle school teachers' assessment and grading practices.* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global: The Humanities and Social Sciences Collection. (UMI No. 305268025)

Grindberg, E. (2014, April 7). Ditching letter grades for a ‘window’ into the classroom. *Cable News Network.* Retrieved from: http://www.cnn.com/2014/04/07/living/ report-card- changes-standards-based-grading-schools/

Guskey, T. R. (1985). *Implementing mastery learning*. Belmont, CA: Wadsworth.

Guskey, T. R. (2000). Grading policies that work against standards...and how to fix them. *IASSP Bulletin, 84*(620), 20–29. doi:10.1177/019263650008462003

Guskey, T. R. (2002, April). *Perspectives on grading and reporting: Differences among teachers, students, and parents.* Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA

Guskey, T. R. (2004). The communication challenge of standards-based reporting. *Phi Delta Kappan, 86*(4), 326–329. doi:10.1177/003172170408600419

Guskey, T. R. (2006). Making high school grades meaningful. *Phi Delta Kappan, 87*(9), 670-675. doi:10.1177/003172170608700910

Guskey, T. R. (2009). Grading policies that work against standards… And how to fix them. In T. R. Guskey (Ed.), *Practical solutions for serious problems in standards-based grading* (pp. 9–26). Thousand Oaks, CA: Corwin.

Guskey, T. R. & Bailey, J. (2001). *Developing grading and reporting systems for student learning*. Thousand Oaks, CA: Corwin.

Guskey, T. R. & Bailey, J. M. (2010). *Developing standards-based report cards.* Thousand Oaks, CA: Corwin.

Guskey, T. R., Swan, G. M. & Jung, L. A. (2010, April). *Developing a statewide, standards-based student report card: A review of the Kentucky initiative*. Paper presented at the Annual Meeting of the American Educational Research Association, Denver, CO.

Hargis, C. H. (1990). *Grades and grading practices: Obstacles to improving education and helping at-risk students.* Springfield, MA: Charles C. Thomas

Hay, P. J., & Macdonald, D. (2008). (Mis)appropriations of criteria and standards-referenced assessment in a performance-based subject. *Assessment in Education: Principles, Policy & Practice, 15*(2), 153–168. doi:10.1080/09695940802164184

Healy, K. L. (1935). A study of the factors involved in the rating of pupils' compositions. *Journal of Experimental Education, 4*(18), 50–53. doi:10.1080/00220973.1935.11009995

Heckman, J. J., & Rubinstein, Y. (2001). The importance of non-cognitive skills: Lessons from the GED testing program. *The American Economic Review, 91*(2), 145– 149. doi:10.2307/2677749

Hill, G. (1935). The report card in present practice. *Educational Method, 15*(2), 115–131. doi:10.1002/j.2333-8504.1973.tb00656.x

Howley, A., Kusimo, P. S. & Parrott, L. (1999). Grading and the ethos of effort. *Learning Environments Research, 3*(3), 229–246. doi:10.1023/A:1011469327430

Hulten, C. E. (1925). The personal element in teachers' marks. *Journal of Educational Research, 12*(1), 49–55. doi:10.1080/00220671.1925.10879575

Imperial, P. (2011). *Grading and reporting purposes and practices in catholic secondary schools and grades' efficacy in accurately communicating student learning* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global: The Humanities and Social Sciences Collection. (UMI No. 896956719)

Jacoby, H. (1910). Note on the marking system in the astronomical course at Columbia College, 1909–1910. *Science, 31*(804), 819–820. doi:10.1126/science.31.804.819

Jimerson, S. R., Egeland, B., Sroufe, L. A., & Carlson, B. (2000). A prospective longitudinal study of high school dropouts examining multiple predictors across development. *Journal of School Psychology, 38*(6), 525– 549. doi:10.1016/S0022-4405(00)00051-0

Jones, A. M. (2008). *High school factors that influence ACT test scores* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global: The Humanities and Social Sciences Collection. (UMI No. 13853)

Kansas State Department of Education. (2015). *Kansas graduation requirements.* Retrieved from http://www.ksde.org/Portals/0/TLA/Graduation%20and%20School%20Choice/Graduation%20and%20Dropout/Kansas%20Graduation%20Requirements%20Fact%20Sheet%202015-2016.pdf

Kansas State Department of Education. (2017a). *KSDE report card – D0*[redacted]. Retrieved from http://ksreportcard.ksde.org/act_scores.aspx?org_no=D0[redacted]&rptType=2

Kansas State Department of Education. (2017b). *KSDE report card – D0*[redacted]. Retrieved from http://ksreportcard.ksde.org/act_scores.aspx?org_no=D0[redacted]&rptType=2

Kansas State Department of Education. (2017c). *KSDE report card – D0*[redacted]. Retrieved from http://ksreportcard.ksde.org/act_scores.aspx?org_no=D0[redacted]&rptType=2

Kelly, F. J. (1914). *Teachers' marks: Their variability and standardization*. New York, NY: Teachers College, Columbia University.

Kelly, T. F. (2008). Grade inflation: Sense and nonsense. *Phi Delta Kappan, 90*(9), 696. doi:10.1177/003172170909000921

Kirschenbaum, H., Simon, S. B., & Napier, R. W. (1971). *Wad-ja-get: The grading game in American education.* New York, NY: Hart.

Klapp Lekholm, A. (2011). Effects of school characteristics on grade in compulsory school. *Scandinavian Journal of Educational Research, 55*(3), 587-608. doi:10.1080/00313831.2011.555923

Klapp Lekholm, A., & Cliffordson, C. (2008). Discrepancies between school grades and tests scores at individual and school level: Effects of gender and family background. *Educational Research and Evaluation, 14*(2), 181–199. doi:10.1080/13803610801956663

Klapp Lekholm, A., & Cliffordson, C. (2009). Effects of student characteristics on grades in compulsory school. *Educational Research and Evaluation, 15*(2), 1–23. doi:10.1080/13803610802470425

Kohn, A. (1999). *Punished by rewards: The trouble with gold stars, incentive plans, As, praise, and other bribes* (2nd ed.). New York, NY: Houghton Mifflin.

Kohn, A. (2002). The dangerous myth of grade inflation. *Chronicle of Higher Education, 49*(11), B7. Retrieved from http://web.b.ebscohost.com/ehost/detail/detail?vid=24&sid=cb1f71f9-7ac6-47c1-98204ad53a0288ed%40sessionmgr120&hid=125&bdata=JnNpdGU9ZWhvc3QtbGl2ZQ%3d%3d#AN=EJ657735&db=eric

Kulick, G., & Wright, R. (2008). The impact of grading on the curve: A simulation analysis. *International Journal for the Scholarship of Teaching and Learning, 2*(2), 5. doi:10.20429/ijsotl.2008.020205

Lauterbach, C. E. (1928). Some factors affecting teachers' marks. *Journal of Educational Psychology, 19*(4), 266– 271. doi:10.1080/14759390000200096

Levin, H. M. (2013). The utility and need for incorporating non-cognitive skills into large-scale educational assessments. In M. von Davier, E. Gonzalez, I. Kirsch, & K. Yamamoto (Eds.), *The role of international large-scale assessments: Perspectives from technology, economy, and educational research* (pp. 67–86). Dordrecht, Netherlands: Springer

Lewis, H. R. (2006). *Excellence without a soul: How a great university forgot education*. New York, NY: Public Affairs.

Linn, R. L. (1982). Ability testing: Individual differences, prediction, and differential prediction. In A. K. Wigdor & W. R. Garner (Eds.), *Ability testing: Uses, consequences, and controversies* (pp. 335–388). Washington, DC: National Academy Press.

Liu, X. (2008, October). *Measuring teachers' perceptions of grading practices: Does school level make a difference?* Paper presented at the Annual Meeting of the Northeastern Educational Research Association, Rocky Hill, CT.

Lloyd, D. N. (1974). Analysis of sixth grade characteristics predicting high school dropout or graduation. *JSAS Catalog of Selected Documents in Psychology, 4*(1), 90. doi:10.1080/00220670903382970

Lloyd, D. N. (1978). Prediction of school failure from third-grade data. *Educational and Psychological Measurement, 38*(4), 1193–1200. doi:10.1177/001316447803800442

Love, D. A., & Kotchen, M. J. (2010). Grades, course evaluations, and academic incentives. *Eastern Economic Journal, 36*(2), 151–163. doi:10.1057/eej.2009.6

Lunenburg, F. C., & Irby, B. J. (2008). *Writing a successful thesis or dissertation: Tips and strategies for students in the social and behavioral sciences*. Thousand Oaks, CA: Corwin Press.

Mahlum, T. (2015). *Exploring the relative value of ACT and GPA of Wyoming high school graduates as a predictor of college success* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global: The Humanities and Social Sciences Collection. (UMI No. 3711733)

Marigold Public Schools. (2017). *Our community*. Retrieved from https://www.usd[redacted]. com/departments/our-community

Marzano, R. J. (1996). Eight questions about implementing standards-based education [Electronic version]. *Practical Assessment, Research, and Evaluation, 5*(6), 2-9. Retrieved from http://files.eric.ed.gov/fulltext/ED410230.pdf

Marzano, R. J. & Heflebower, T. (2011). Grades that show what students know. *Educational Leadership, 69*(3), 34–39. Retrieved from https://eric.ed.gov/?id =EJ963096

McCandless, B. R., Roberts, A., & Starnes, T. (1972). Teachers' marks, achievement test scores, and aptitude relations with respect to social class, race, and sex. *Journal of Educational Psychology, 63*(2), 153–159. doi:10.1037/h0032646

McCann, M. D. (2003). *Grade inflation: The relationship between mathematics report card grades and standardized test results at the elementary level* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global: The Humanities and Social Sciences Collection. (UMI No. 3117815)

McDaniel. D. (2015). *ACT scores show students not ready for college coursework.* Retrieved from http://www.decaturdaily.com/news/lawrence_county/act-scores-show-students-not-ready-for-college-coursework/article_babb3143-5cd3-57bd-aeaf-925d3679b08b.html

McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice, 20*(1), 20– 32. doi:10.1111/j.1745-3992.2001.tb00055.x

McMillan, J. H. (2009). Synthesis of issues and implications for practice. In T.R. Guskey (Ed.), *Practical solutions for serious problems in standards-based grading* (pp. 105–120). Thousand Oaks, CA: Corwin Press.

McMillan, J. H., & Lawson, S. R. (2001). *Secondary science teachers' classroom assessment and grading practices.* Retrieved from ERIC database (ED 450 158).

McMillan, J. H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *Journal of Educational Research, 95*(4), 203–213. doi:10.1080/00220670209596593

McMillan, J. H., & Nash, S. (2000, April). *Teacher classroom assessment and grading decision making.* Paper presented at the Annual Meeting of the National Council of Measurement in Education, New Orleans, LA.

McMunn, N., Schenck, P., & McColskey, W. (2003, April). *Standards-based assessment, grading, and reporting in classrooms: Can district training and support change teacher practice?* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

Melograno, V. J. (2007). Grading and report cards for standards-based physical education. *Journal of Physical Education, Recreation, and Dance, 78*(6), 45–53. doi:10.1080/07303084.2007.10598041

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed.) (pp. 13– 103). New York, NY: American Council of Education and Macmillan.

Meyer, M. (1908). The grading of students. *Science, 28*(712), 243–252. doi:10.1126/science.28.712.243

Miner, B. C. (1967). Three factors of school achievement. *The Journal of Educational Research, 60*(8), 370–376. doi:10.2307/27531890

Mohnsen, B. (2013). Assessment and grading in physical education. *Strategies: A journal for physical and sports educators, 20*(2), 24–28. doi:10.1080/08924562.2006.10590709

Moore, C. C. (1939). The elementary school mark. *The Pedagogical Seminary and Journal of Genetic Psychology, 54*(2), 285–294. doi:10.1080/08856559.1939.10534336

Morris, J. D., Ehren, B. J., & Lenz, B. K. (1991). Building a model to predict which fourth through eighth-graders will drop out in high school. *Journal of Experimental Education, 59*(3), 286–293. doi:10.1080/00220973.1991.10806615

Myford, C. (2012). Rater cognition research: Some possible directions for the future. *Educational Measurement: Issues and Practice, 31*(3), 48–49. doi:10.1111/j.1745-3992.2012.00243.x

Nichols, S. L., & Berliner, D. C. (2007). *Collateral damage: How high stakes testing corrupts America's schools.* Cambridge, MA: Harvard Education Press.

Nicolson, F. W. (1917). Standardizing the marking system. *Educational Review, 54*(1), 225–237. doi:10.1037/h0069798

Noble, J., Davenport, M., & Sawyer, R. (2001). Relationships between non-cognitive characteristics, high school course work and grades, and performance on a college admissions test. Paper presented at the annual meeting of the American Education Research Association in Seattle, WA.

Noble, J. P. & Schnelker, D. (2007). *Using hierarchical modeling to examine course work and ACT score relationships across high schools.* (Research Report 2007-2). Iowa City, IA: ACT. Retrieved from http://files.eric.ed.gov/fulltext/ED510476.pdf

O'Connor, K. (2009). *How to grade for learning: Linking grades to standards.* (3rd ed.). Glenview, IL: Pearson.

Office of Institutional Research and Policy Studies (OIRP). (2011). *High school GPA and SAT score effects on graduation rates*. Boston, MA: Boston University Press.

O'Halloran, K. C., & Gordon M. E. (2014). A synergistic approach to turning the tide of grade inflation. *Journal of Higher Education, 68*(1), 1005-1023. doi:10.1007/s10734-014-9758-5

Olsen, D. R. (1995). *Grade inflation: Reality or Myth? Student preparation level vs. grades at Brigham Young University 1975-1994* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global: The Humanities and Social Sciences Collection. (UMI No. 9608616)

Pallas, A. M. (2003). Educational transitions, trajectories, and pathways. In J. T. Mortimer & M. J. Shanahan (Eds.), *Handbook of the life course*. New York, NY: Kluwer Academic/Plenum.

Pattison, E., Grodsky, E., & Muller, C. (2013). Is the sky falling? Grade inflation and the signaling power of grades. *Educational Researcher, 42*(5), 259–265. doi:10.3102/0013189x13481382

Pearson, K. (1930). *Life of Francis Galton.* London, UK: Cambridge University Press.

Polikoff, M. S., Porter, A. C., & Smithson, J. (2011). How well aligned are state assessments of student achievement with state content standards? *American Educational Research Journal, 48*(4), 965–995. doi:10.3102/0002831211410684

Pope, D. C. (2011). *"Doing school": How we are creating a generation of stressed out, materialistic, and mis-educated students*. New Haven, CT: Yale University Press.

Quann, C. J. (1983). *Grades and grading: Historical perspectives and the 1982 AACRAO study*. Washington, DC: American Association of Collegiate Registrars and Admissions Officers.

Randall, J., & Engelhard, G. (2009). Examining teacher grades using Rasch measurement theory. *Journal of Educational Measurement, 46*(1), 1–18. doi:10.1111/j.1745-3984.2009.01066.x

Randall, J., & Engelhard, G. (2010). Examining the grading practices of teachers. *Teaching and Teacher Education*, *26*, 1372–1380. doi:10.1016/j.tate.2010.03.008

Resnick, L. B. (1987). The 1987 presidential address: Learning in school and out. *Educational Researcher*, *16*(9), 13–20 + 54. doi:10.3102/0013189X016009013

Roberts, W. L., & Noble, J. P. (2004). *Academic and non-cognitive variables related to ACT Plan scores* (ACT Research Report Series 2004-1). Retrieved on January 1 2018 from http://www.act.org/research/reports/index.html

Roderick, M., & Camburn, E. (1999). Risk and recovery from course failure in the early years of High School. *American Educational Research Journal, 36*(2), 303– 343. doi:10.3102/00028312036002303

Rojstaczer, S., & Healy, C. (2012). Where A is ordinary: The evolution of American college and university grading, 1940–2009. *Teachers College Record, 114*(7), 1–23. Retrieved from https://eric.ed.gov/?id=EJ1001967

Ross, C. C., & Hooks, N. T. (1930). How shall we predict high-school achievement? *The Journal of Educational Research, 22*(3), 184–196. doi:10.2307/27525222

Ross, J. A., & Kostuch, L. (2011). Consistency of report card grades and external assessments in a Canadian province. *Educational Assessment, Evaluation, and accountability, 23*(2), 158–180. doi:10.1007/s11092-011-9117-3

Roth, J., Crans, G., Carter, R. L., Ariet, M., & Resnick, M. B. (2000). Effect of high school course-taking and grades on passing a college placement test, *High School Journal, 84*(2), 72-87. Retrieved from https://eric.ed.gov/?id=ED428817

Rugg, H. O. (1918). Teachers' marks and the reconstruction of the marking system. *The Elementary School Journal, 18*(9), 701–719. doi:10.1086/454643

Rumberger, R. W. (1987). High school dropouts: A review of issues and evidence. *Review of Educational Research, 57*(2), 101–121. doi:10.3102/00346543057002101

Rumberger, R. W. (2011). *Dropping out: Why students drop out of high school and what can be done about it.* Cambridge, MA: Harvard University Press.

Russell, J. A., & Austin, J. R. (2010). Assessment practices of secondary music teachers. *Journal of Research in Music Education, 58*(1), 37–54. doi:10.1177/0022429409360062

Sawyer, R. (2013). Beyond correlations: Usefulness of high school GPA and test scores in making college admissions decisions. *Applied Measurement in Education, 26*(2), 89-112. doi:10.1080/08957347.2013.765433

Schiel, J., Pommerich, M., & Noble, J. P. (1996). *Factors associated with longitudinal educational achievement, as measured by ACT Plan and ACT assessment scores. ACT Research Reports Series*. Retrieved from http://www.act.org/research/researchers/reports/pdf/ACT_RR96-05.pdf

Schmidt, P. (2007, March 9). High-school students aim higher without learning more, federal studies find. *Chronicle of Higher Education, 53*(27), A32. Retrieved from http://web.b.ebscohost.com/ehost/detail/detail?vid=22&sid=cb1f71f9-7ac6-47c1-98204ad53a0288ed%40sessionmgr120&hid=125&bdata=JnNpdGU9ZWhvc3QtbGl2ZQ%3d%3d#AN=EJ757229&db=eric

Schneider, J., & Hutt, E. (2014). Making the grade: A history of the A-F marking scheme. *Journal of Curriculum Studies, 46*(2), 201–224. doi:10.1080/00220272.2013.790480.

Scriffiny, P. L. (2008). Seven reasons for standards-based grading. *Educational Leadership, 66*(2), 70–74. Retrieved from http://www.ascd.org/publications/educational_leadership/oct08/vol66/num02/Seven_Reasons_for_StandardsBased_Grading.aspx

Shippy, N., Washer, B. A., & Perrin, B. (2013). Teaching with the end in mind: The role of standards-based grading. *Journal of Family and Consumer Sciences, 105*(2), 14–16. doi: 10.14307/JFCS105.2.5

Shriner, W. O. (1930). The comparison factor in the evaluation of examination papers. *Teachers College Journal, 1*(1), 65–74. doi:10.3102/0034654316672069

Silberstein, N. (1922). The variability of teachers' marks. *English Journal, 11*(7), 414–424. Retrieved from https://www.jstor.org/stable/802757

Sims, V. M. (1933). Reducing the variability of essay examination marks through eliminating variations in standards of grading. *Journal of Educational Research, 26*(9), 637–647. doi:10.1080/00220671.1933.10880358

Simon, S. B., & Bellanca, J. A. (1976). *Degrading the grading myths: A primer of alternatives to grades and marks.* Washington, DC: Association for Supervision and Curriculum Development.

Simon, M., Tierney, R. D., Forgette-Giroux, R., Charland, J., Noonan, B., & Duncan, R. (2010). A secondary school teacher’s description of the process of determining report card grades. *McGill Journal of Education, 45*(3), 535–554. doi:10.7202/1003576ar

Smith, A. Z., & Dobbin, J. E. (1960). Marks and marking systems. In C. W. Harris (Ed.), *Encyclopedia of educational research* (3rd ed.) (pp. 783–791). New York, NY: Macmillan.

Stanley, G. & Baines, L. (2004). No more shopping for grades at B-Mart: Re-establishing grades as indicators of academic performance. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 77*(3), 101-104. doi:10.1080/00098650409601237

Starch, D. (1913). Reliability and distribution of grades. *Science, 38*(983), 630–636. doi:10.1126/science.38.983.630

Starch, D. (1915). Can the variability of marks be reduced? *School and Society, 2*(1), 242–243. Retrieved from http://jstor.org/stable/994119

Starch, C., & Elliott, E. C. (1912). Reliability of the grading of high school work in English. *School Review, 20*(1). Retrieved from http://www.jstor.org/stable/ 1076706

Starch, C., & Elliott, E. C. (1913). Reliability of grading work in mathematics. *School Review, 21*(1). Retrieved from http://www.jstor.org/stable/1076246

Stumpf, H., & Stanley, J. C. (2002). Group data on high school grade point average and scores on academic aptitude tests as predictors of institutional graduation rates. *Educational and Psychological Measurement, 62*(6), 1042-1052. doi:10.1177/0013164402238091

Sun, Y., & Cheng, L. (2013). Teachers' grading practices: Meaning and values assigned. *Assessment in Education: Principles, Policy & Practice, 21*(3), 326– 343. doi:10.1080/0969594.2013.768207

Sunflower Public Schools. (2017). *2017 district snapshot*. Retrieved from https://www.usd[redacted].org/domain/954

Svennberg, L., Meckbach, J., & Redelius, K. (2014). Exploring PE teachers' 'gut feelings': An attempt to verbalise and discuss teachers' internalised grading criteria. *European Physical Education Review, 20*(2), 199–214. doi:10.1177/1356336X13517437

Swan, G. M., Guskey, T. R., & Jung, L. A. (2014). Parents' and teachers' perceptions of standards-based and traditional report cards. *Educational Assessment, Evaluation and Accountability, 26*(3), 289–299. doi:10.1007/s11092-014-9191-4

Swineford, F. (1947). Examination of the purported unreliability of teachers' marks. *The Elementary School Journal*, *47*(9), 516–521. doi:10.2307/3203007

Thorsen, C., & Cliffordson, C. (2012). Teachers' grade assignment and the predictive validity of criterion-referenced grades. *Educational Research and Evaluation, 18*(2), 153–172. doi:10.1080/13803611.2012.659929

Tierney, R. D., Simon, M., & Charland, J. (2011). Being fair: Teachers' interpretations of principles for standards-based grading. *The Educational Forum, 75*(3), 210–227. doi:10.1080/00131725.2011.577669

Troob, C. (1985). *Longitudinal study of students entering high school in 1979: The relationship between first term performance and school completion.* New York, NY: New York City Board of Education.

Troug, A. J., & Friedman, S. J. (1996). *Evaluating high school teachers' written grading policies from a measurement perspective.* Paper presented at the annual meeting of the National Council on Measurement in Education, New York City, NY.

Unzicker, S. P. (1925). Teachers' marks and intelligence. *The Journal of Educational Research, 11*(2), 123–131. doi:10.1080/00220671.1925.10879537

U.S. Department of Education. (2012a). *Science literacy: Average scores (overall).* Retrieved from https://nces.ed.gov/surveys/pisa/pisa2012/pisa2012highlights_4a.asp

U.S. Department of Education. (2012b). *Mathematics literacy: Average scores (overall).* Retrieved from https://nces.ed.gov/surveys/pisa/pisa2012/pisa2012highlights_3a.asp

U.S. Department of Education. (2012c). *Reading literacy: Average scores (overall).* Retrieved from https://nces.ed.gov/surveys/pisa/pisa2012/pisa2012highlights_5a.asp

Varlas, L. (2013). How we got grading wrong, and what to do about it. *Education Update*, *55*(10), 6-7. Retrieved from http://www.ascd.org/publications/newsletters/education-update/oct13/vol55/num10/How-We-Got-Grading-Wrong,-and-What-to-Do-About-It.aspx

Vatterott, C. (2015). *Rethinking grading: Meaningful assessment for standards-based learning.* Alexandria, VA: Association for Supervision and Curriculum Development.

Voss, H. L., Wendling, A., & Elliott, D. S. (1966). Some types of high school dropouts. *The Journal of Educational Research*, *59*(8), 363–368. doi:10.1080/00220671.196 6.10883382

Walsh, P. (2010). Does competition among schools encourage grade inflation? *Journal of School Choice, 4*(1), 149-173. doi:10.1080/15582159.2010.483918

Welsh, M. E., & D'Agostino, J. (2009). Fostering consistency between standards-based grades and largescale assessment results. In T. R. Guskey (Ed.), *Practical solutions for serious problems in standards-based grading* (pp. 75–104). Thousand Oaks, CA: Corwin Press.

Welsh, M. E., D'Agostino, J. V., & Kaniskan, R. (2013). Grading as a reform effort: Do standards-based grades converge with test scores? *Educational Measurement: Issues and Practice, 32*(2), 26–36. doi:10.1111/emip.12009

Westrick, P. A., Le, H., Robbins, S. B., Radunzel, J. M. R., & Schmidt, F. L. (2015). College performance and retention: A meta-analysis of the predictive validities of ACT scores, high school grades, and SES. *Educational Assessment, 20*(1), 23-45. doi:10.1080/10627197.2015.997614

Wiggins, G. (1994). Toward better report cards. *Educational Leadership, 52*(2), 28-37. Retrieved from http://ascd.org/publication/educationalleadership/oct94/vol52/num02/Toward-Better-ReportCards.aspx

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. doi:10.1002/j.2333-8504.2000.tb01838.x

Winter, R. (1993). Education or grading? Arguments for a non-subdivided honors degree. *Studies in Higher Education, 18*(3), 363-377. doi:10.1080/03075079312331382271

Woodruff, D. J., & Ziomek, R. L. (2004). *High school grade inflation from 1991 to 2003*. Iowa City, IA: ACT.

Yesbeck, D. M. (2011). *Grading practices: Teachers' considerations of academic and non-academic factors* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global: The Humanities and Social Sciences Collection. (UMI No. 913076079)

Zellner, K. M. (2008). *An examination of grade distribution patterns and grade inflation at a baccalaureate nursing program* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses Global: The Humanities and Social Sciences Collection. (UMI No. 3304475)

Zhang, Q. & Sanchez, E. I. (2013). High school grade inflation from 2004 to 2011. *ACT Report Series, 6*(1), 1-18. Retrieved from http://files.eric.ed.gov/fulltext/ ED555577.pdf

Zirkel, P. A. (2007) Grade inflation: High schools' skeleton in the closet. *Education Week, 26*(29), 30. Retrieved from http://www.edweek.org/ew/articles /2007/03/27/29zirkel.h26.html

Appendices

Appendix A: Sunflower Public Schools Research Request Form

[redacted] Public Schools, [redacted]

RESEARCH PROPOSAL

Investigator(s): Loren Scarbrough | Date: 11-16-16

Mailing Address: 10118 E. Funston St. Wichita, KS 67207 | Telephone: 316-258-7644

Email Address: lscarbrough@usd259.net

Name and Address of Company, University/College, School/Department: Baker University, 8001 College Boulevard, Suite 100, Overland Park, KS 66210

University/College Advisor (applicable to students only): Dr. Sharon Zoellner

---

Complete this form using brief, concise statements and send one copy to the Department of Innovation & Evaluation, [redacted] Public Schools, [redacted] for presentation to the Research Council (or email to [redacted]. This form must be dated and signed by a majority of the [redacted] Research Council members before commencement of any new research project. The investigator(s) agrees upon completion of the research project to submit a copy of the final report to Innovation & Evaluation, [redacted]

1. Title or brief description of the proposed study:

   Cumulative ACT Scores and Core High School GPAs: The Relationship and Influential Factors in South Central Kansas

2. Statement of the educational problem:

   The landscape of American education has changed drastically since the 19th century, but the original grading paradigm is still prevalent in today's educational practices (Vatterott, 2015). Grading has been at the heart of the kindergarten through post-secondary experience for 120 years. Regardless of the original validity of criteria used, grades have come to epitomize the academic ability, which leaves teachers, students, parents, and college admissions staff to question whether a grade is truly indicative of mastery of standards, particularly for at-risk youth (Pope, 2011). As recently as 2012, reports released by the United States Department of Education (USDE) conveyed a paradox in student GPAs in science, mathematics, and reading courses and corresponding scores on federal standardized tests (USDE, 2012a, 2012b, 2012c). While student enrollment and GPAs in upper level courses across the United States are increasing, the U.S. has consistently scored lower than other similarly developed countries on international standardized tests (Schmidt, 2007). The divergence between GPAs and standardized test scores is widening for students with a motivation or cognitive skill discrepancy according to the Office of Institutional Research and Policy Studies (OIRP, 2011). While extensive research has been conducted on the predictive ability of high school GPAs and ACT scores for college readiness (Beecher & Fisher, 1999; Caruth & Caruth, 2013; Coyle & Pillow, 2008; Geiser & Santelices, 2007; Sawyer, 2013; Stumpf & Stanley, 2002; Westrick, Le, Robbins, Radunzel, & Schmidt, 2015; Zellner, 2008), there is minimal research on the relationship between cumulative ACT scores and high school GPAs (Woodruff & Ziomek, 2004; Zhang & Sanchez, 2013).

3. Specific purpose and expected outcomes:

   This study will be conducted to examine the relationship between composite ACT scores and GPAs in the core courses of English, mathematics, history, and science in grades nine through 12 in high schools in south central Kansas. An additional purpose of the study will be to explore how gender, race, socioeconomic status, and high school categorization (urban, suburban, or rural) affect the relationship between composite ACT scores and GPAs in the core courses of English, mathematics, history, and science in grades nine through 12 in high schools in south central Kansas.

[redacted] and [redacted] have already agreed to provide data to participate in this study as long as they are provided a copy of the results. The results of this study could foster intra- and inter-district dialogue about grading practices, standardized test prep, Common Core content standards, and bridging the skills or knowledge gap should one be shown to exist. If there is not a relationship between core GPA and composite ACT score, school districts could discuss and further analyze why there is not a relationship and why core course grades are not reflective of core course skills as measured by the ACT.

4. Hypothesis(es) to be tested (if applicable):

   RQ1. To what extent is there a relationship between cumulative ACT scores and core GPAs (English, mathematics, history, and science) for students in grades nine through 12 in the south central Kansas districts?

   H1. There is a relationship between cumulative ACT scores and core GPAs (English, mathematics, history, and science) for students in grades nine through 12 in the south central Kansas districts.

   RQ2. To what extent is the relationship between cumulative ACT scores and

   core GPA (English, mathematics, history, and science) for students in grades nine through 12 in the south central Kansas districts affected by gender, race, socioeconomic status, and community type (urban, suburban, or rural)?

   H2. There is a difference in the relationship between cumulative ACT scores and core GPAs (English, mathematics, history, and science) for students in grades nine through 12 in the south central Kansas districts between genders.

   H3. There is a difference in the relationship between cumulative ACT scores and core GPAs (English, mathematics, history, and science) for students in grades nine through 12 in the south central Kansas districts among different races.

   H4. There is a difference in the relationship between cumulative ACT scores and core GPAs (English, mathematics, history, and science) for students in grades nine through 12 in the south central Kansas districts between SES levels.

   H5. There is a difference in the relationship between cumulative ACT scores and core GPAs (English, mathematics, history, and science) for students in grades nine through 12 in the south central Kansas districts between community types.

5. What specific [redacted] Strategic Plan Objectives have you identified as being directly related to this proposal? State the relationship (see enclosed listing) and how [redacted] will benefit from this research:

   Objective 3: The academic skill and knowledge gap among the student populations will be continually reduced until eliminated as measured by multiple assessments.

Objectice 5: A coherent, rigorous, safe and nurturing, culturally responsive and inclusive learning community will be fostered and sustained.

Similar to Geiser and Santelices' (2007) study on the predictive ability of high school grades for college success, the results of this study may foster a review of effective grading practices, encourage the expansion of professional development models, and promote conversation regarding the validity of using GPAs and standardized test scores to measure grade inflation. It is progressively more important for principals and superintendents to evaluate the relationship between student grades and standardized test scores for their schools and districts to better support accurate assessment of the mastery of core content standards (Beecher & Fisher, 1999). Additionally, this study may allow schools to identify indicators of standardized test failure and respond more quickly to student needs that are masked by GPAs. High school administrators, staff, parents, and students all benefit from the results by understanding how grading practices and demographic information impact the need for college readiness remediation as measured by ACT scores. Finally, Woodruff and Ziomek (2004) and Zhang and Sanchez (2013) published the two most prominent studies on the relationship between cumulative ACT scores and GPAs, but the studies had divergent results. The results of this study contribute to the overall body of knowledge where negligible research has been conducted and the results have been conflicting.

6. Description of sample needed: grade levels, students, teachers, and/or management employees, and desired location(s) if there is a preference.

   The data set needed is the gender, race, SES status, core GPA, and composite ACT score of all 2017 graduates of [redacted] with an ACT score on file.

7. Procedures and methods to be employed. (What will be done by the investigator and participants in the study, data to be gathered, and data gathering instruments to be used?) If possible, attach copy of instrument(s) to be used in gathering data.

   Participants will not have an active role in this study, since only archival data is needed. Contact has been made with [redacted] Executive Director of Innovation and Evaluation to pull the requested data. Data requested (see question 7) will be input into an Excel spreadsheet (see attached file).

8. Data treatment and analysis:

   The data will not include identification markers, will be stored on a password protected external hard drive, and will be deleted after the completion of the dissertation and publishing process. Analysis will include Pearson product moment correlation coefficients to determine the strength and direction of the relationship between core GPAs and composite ACT scores and Fisher's z tests to determine the differences in relationships for each sub-group.

9. Expected starting date: January 2018

   Duration of study: six months

   Expected completion date of dissertation or final report: June 2018

10. Protection of human subjects:

    a. Rights of privacy guaranteed — Yes ☒ No ☐
    b. Permission for participation on record — Yes ☒ No ☐
    c. Clearance by company, university/college/school — Yes ☒ No ☐

Appendix B: Sunflower Public Schools Research Approval Letter

December 5, 2016

Loren Scarbrough
10118 E. Funston St.
Wichita, Kansas 67207

Re: Research Proposal dated 11-16-16

Dear Ms. Scarbrough,

This letter is in response to your recent research request *Cumulative ACT Scores and Core High School GPAs: The Relationship and Influential Factors in South Central Kansas* in the [redacted] Public Schools. At this time I am pleased to inform you that the Research Council has given approval for your study.

As you proceed with your study, please note that this letter approves the research project as described above, but that it is incumbent upon the researcher(s) to negotiate distribution. The project also must not unduly increase the workload of any employee of the [redacted] Public Schools. The [redacted] Public Schools staff has the right to discontinue participation at any time. If for any reason it becomes necessary to modify what was originally presented in your proposal, the Research Council must be so informed and approve any changes in advance. Copies of any reports related to this research must be submitted to the office of Assessments and Research and be made available to the participating schools as well.

Please direct any changes or questions to [redacted] at [redacted]@usd[redacted].net.

On behalf of the USD [redacted] Research Council,

[redacted]

[redacted] Executive Director
Assessments and Research
[redacted] Public Schools, USD [redacted]

Appendix C: E-mail Requesting Permission from Daisy Public Schools

From: Loren Scarbrough [mailto: lscarbrough@usd259.net]
Sent: Friday, December 09, 2016 12:15 PM
To: [redacted]
Subject: Research in [redacted] Public Schools

Mr. [redacted],

I hope this e-mail finds you well. My name is Loren Scarbrough and I am currently completing my doctoral coursework through Baker University. I am reaching out to you, because I would like to use some data from the [redacted] Public Schools' class of 2017.

This study will be conducted to examine the relationship between composite ACT scores and core GPAs (English, mathematics, history, and science) for 2017 graduates of high schools in south central Kansas, and to explore how gender, race, socioeconomic status, and high school community type (urban, suburban, or rural) affect said relationship. My goal is collect the composite ACT score, core GPA, gender, race, and socioeconomic status of each member of the class of 2017 from districts across [redacted] and [redacted] counties.

Would you district potentially be willing to provide anonymous student data in an Excel file for my dissertation research, and if so, what are the procedures for obtaining data from [redacted] Public Schools?

Thank you for your time and support in this matter.

Loren Scarbrough

Appendix D: E-mail Requesting Permission from Marigold Public Schools

From: Loren Scarbrough [mailto: lscarbrough@usd259.net]
Sent: Friday, December 09, 2016 12:15 PM
To: [redacted]
Subject: Research in [redacted] Public Schools

Dr. [redacted],

I hope this e-mail finds you well. My cohort is advancing rapidly through our coursework and I am getting ever closer to the research phase of my dissertation. I am reaching out to you, because I would like to use some data from the [redacted] Public Schools' class of 2017.

This study will be conducted to examine the relationship between composite ACT scores and core GPAs (English, mathematics, history, and science) for 2017 graduates of high schools in south central Kansas, and to explore how gender, race, socioeconomic status, and high school community type (urban, suburban, or rural) affect said relationship. My goal is collect the composite ACT score, core GPA, gender, race, and socioeconomic status of each member of the class of 2017 from districts across [redacted] and [redacted] counties.

Would you district potentially be willing to provide anonymous student data in an Excel file for my dissertation research, and if so, what are the procedures for obtaining data from [redacted] Public Schools?

Thank you for your time and support in this matter.

Loren Scarbrough

Appendix E: E-mail Granting Permission from Daisy Public Schools

From: [redacted]
Sent: Monday, December 12, 2016 3:02 PM
To: Loren Scarbrough <lscarbrough@usd259.net>
Subject: Re: Research in [redacted] Public Schools

Loren-

I would like to see the results of your research once completed. Be in touch when you need the data.

On Mon, Dec 12, 2016 at 9:30 AM, Loren Scarbrough <lscarbrough@usd259.net> wrote:

[redacted],

Core classes include any courses for mathematics, English, science, or social studies credit.

Loren Scarbrough

From: [redacted]
Sent: Monday, December 12, 2016 8:14 AM
To: Loren Scarbrough <lscarbrough@usd259.net>
Subject: Re: Research in [redacted] Public Schools

Loren-

Can you also tell me what courses you've identified in your research that qualify as "core"?

On Mon, Dec 12, 2016 at 7:27 AM, Loren Scarbrough <lscarbrough@usd259.net> wrote:

[redacted],

Thank you for your prompt reply. My goal is to have all the data by early spring of 2018. I am reaching out now because I know I am asking for a lot. I have attached a rudimentary spreadsheet that can be used for input, should your district choose to participate and wish to use it later on.

Loren Scarbrough

From: [redacted]
Sent: Monday, December 12, 2016 7:10 AM
To: Loren Scarbrough <lscarbrough@usd259.net>
Subject: Re: Research in [redacted] Public Schools

Loren-

We are interested in helping you out. I am currently trying to find out if we can produce this for you and what the timeline might be. Our database administrator is currently bogged-down with another large project. Can you tell me what your deadline for this information would be if we participated?

Thanks.

Appendix F: E-mail Granting Permission from Marigold Public Schools

Appendix G: Baker University IRB Form

BAKER UNIVERSITY

IRB Request

Date 04-02-2018

IRB Protocol Number ____________ (IRB use only)

I. Research Investigator(s) (students must list faculty sponsor)

Department(s) School of Education Graduate Department

	Name	Signature	
1.	Loren Scarbrough	Loren Scarbrough Digitally signed by Loren Scarbrough Date: 2018.04.02 20:42:26 -05'00'	Principal Investigator
2.	Sharon Zoellner		☑ Check if faculty sponsor
3.	Peg Waterman		☐ Check if faculty sponsor
4.			☐ Check if faculty sponsor

Principal investigator contact information

Phone 316-258-7644

Email LorenBScarbrough@stu.bakeru.edu

Address 10118 E. Funston St.
Wichita, KS 67207

Note: When submitting your finalized, signed form to the IRB, please ensure that you cc all investigators and faculty sponsors using their official Baker University (or respective organization's) email addresses.

Faculty sponsor contact information

Phone 913-344-1225

Email Sharon.zoellner@bakeru.edu

Expected Category of Review: ☑ Exempt ☐ Expedited ☐ Full ☐ Renewal

II. Protocol Title

The Relationship Between Composite ACT Scores and Core GPAs for South-Central Kansas High School Graduates in the Class of 2017

Baker IRB Submission form page 1 of 4

III. Summary:

The following questions must be answered. Be specific about exactly what participants will experience and about the protections that have been included to safeguard participants from harm.

A. In a sentence or two, please describe the background and purpose of the research.

This study will be conducted to examine the relationship between composite ACT scores and core GPAs (English, mathematics, history, and science) in grades nine through 12 in high schools in south central Kansas. An additional purpose of the study is to explore how gender, race, socioeconomic status, and high school community type (urban, suburban, or rural) affect the relationship between composite ACT scores and core GPAs (English, mathematics, history, and science) in grades nine through 12 in high schools in south-central Kansas.
Three school districts from south central Kansas have agreed to participate in this study: [redacted] Public Schools, [redacted] Public Schools, and [redacted] Public Schools. [redacted] Public Schools is an urban district located in

B. Briefly describe each condition, manipulation, or archival data set to be included within the study.

There are no conditions or manipulations included within this study.

Archival data regarding student gender, race, socioeconomic status, core high school GPA and composite ACT score will be used in this study. Student gender (male or female) will be indicated on each student's personal information form. Student race will be indicated on each student's school record. The following races will be included in this study: Black, White, Hispanic, Asian, and other. For the purpose of the data analysis students categorized as Native American, Alaskan Native, Native Hawaiian, other Pacific Islander, or Multiracial will be

IV. Protocol Details

A. What measures or observations will be taken in the study? If any questionnaire or other instruments are used, provide a brief description and attach a copy.

Four categorical variables will be used in this study: gender, race, socioeconomic status, and community type. Student gender option are male or female. Student race will be limited to Black, White, Hispanic, Asian, and other. For the purpose of the data analysis students categorized as Native American, Alaskan Native, Native Hawaiian, other Pacific Islander, or Multiracial will be included in the other category. Student socioeconomic status will be free or reduced lunch or non-free or reduced lunch as determined by federal guidelines for parent or guardian income. Community type (urban, suburban, or rural) will be determined by geographic location and U.S. Census Bureau categorization of the school district from which a student graduated. Students from [redacted] Public

B. Will the subjects encounter the risk of psychological, social, physical, or legal risk? If so, please describe the nature of the risk and any measures designed to mitigate that risk.

The subjects will not encounter the risk of psychological, social, physical, or legal risk during this study.

C. Will any stress to subjects be involved? If so, please describe.

Stress to subjects will not be involved in this study.

D. Will the subjects be deceived or misled in any way? If so, include an outline or script of the debriefing.

Subjects will not be deceived or misled in any way during this study.

E. Will there be a request for information which subjects might consider to be personal or sensitive? If so, please include a description.

There will not be a request for information which subjects might consider to be personal or sensitive.

F. Will the subjects be presented with materials which might be considered to be offensive, threatening, or degrading? If so, please describe.

Subjects will not be presented with materials which might be considered to be offensive, threatening, or degrading during this study.

G. Approximately how much time will be demanded of each subject?

Time will not be demanded of any subject during this study.

H. Who will be the subjects in this study? How will they be solicited or contacted? Provide an outline or script of the information which will be provided to subjects prior to their volunteering to participate. Include a copy of any written solicitation as well as an outline of any oral solicitation.

The population for the current study includes students who took the ACT in grades nine through 12 and graduated from a south-central Kansas high school in the spring of 2017. Students with composite ACT scores from 13 high schools in three districts will comprise the population. The sample will consist of students who graduated in the spring of 2017 and took the ACT during the 2015-2016 or 2016-2017 school years. Archival data is being used so subjects will not be solicited or contacted.

I. What steps will be taken to insure that each subject's participation is voluntary? What if any inducements will be offered to the subjects for their participation?

No inducements will be offered to the subjects for their participation.

J. How will you insure that the subjects give their consent prior to participating? Will a written consent form be used? If so, include the form. If not, explain why not.

A written consent form will not be used since archival data will be used in this study.

K. Will any aspect of the data be made a part of any permanent record that can be identified with the subject? If so, please explain the necessity.

No aspect of the data will be made part of any permanent record that can be identified with any subject.

L. Will the fact that a subject did or did not participate in a specific experiment or study be made part of any permanent record available to a supervisor, teacher, or employer? If so, explain.

A subject's participation or non-participation in the study will not be made part of any permanent record available to a supervisor, teacher, or employer.

M. What steps will be taken to insure the confidentiality of the data? Where will it be stored? How long will it be stored? What will be done with the data after the study is completed?

The Excel spreadsheets provided by each of the participating school districts will not include student name or identification number. Data will be stored on a password protected external hard drive for three years. At the end of the three-year time period, the hard drive will be re-imaged and all data deleted permanently.

N. If there are any risks involved in the study, are there any offsetting benefits that might accrue to either the subjects or society?

There are no risks involved in this study.

O. Will any data from files or archival data be used? If so, please describe.

Archival data regarding student gender, race, socioeconomic status, core high school GPA and composite ACT score will be used in this study. Student gender (male or female) will be indicated on each student's personal information form. Student race will be indicated on each student's school record. The following races will be included in this study: Black, White, Hispanic, Asian, and other. For the purpose of the data analysis students categorized as Native American, Alaskan Native, Native Hawaiian, other Pacific Islander, or Multiracial will be included in the other category since those racial groups could have small sample sizes. Student socioeconomic status (free or reduced lunch or non-free or reduced lunch) will be determined by federal guidelines for parent or

Appendix H: Baker University IRB Approval

Baker University Institutional Review Board

April 4th, 2018

Dear Loren Scarbrough and Sharon Zoellner,

The Baker University IRB has reviewed your project application and approved this project under Expedited Status Review. As described, the project complies with all the requirements and policies established by the University for protection of human subjects in research. Unless renewed, approval lapses one year after approval date.

Please be aware of the following:

1. Any significant change in the research protocol as described should be reviewed by this Committee prior to altering the project.
2. Notify the IRB about any new investigators not named in original application.
3. When signed consent documents are required, the primary investigator must retain the signed consent documents of the research activity.
4. If this is a funded project, keep a copy of this approval letter with your proposal/grant file.
5. If the results of the research are used to prepare papers for publication or oral presentation at professional conferences, manuscripts or abstracts are requested for IRB as part of the project record.

Please inform this Committee or myself when this project is terminated or completed. As noted above, you must also provide IRB with an annual status report and receive approval for maintaining your status. If you have any questions, please contact me at npoell@bakeru.edu or 785.594.4582.

Sincerely,

Nathan D. Poell

Nathan Poell, MA
Chair, Baker University IRB

Baker University IRB Committee
Scott Crenshaw
Jamin Perry, PhD
Susan Rogers, PhD
Joe Watson, PhD