January 2016

School Accountability Grades: The Influence Of Student And School Characteristics

Tara Isaacs
Eastern Kentucky University

SCHOOL ACCOUNTABILITY GRADES: THE INFLUENCE OF STUDENT AND SCHOOL CHARACTERISTICS

By

Tara Grimes Isaacs

Dissertation Approved:

Dr. Charles Hausman, Advisory Committee Chair

Dr. James Bliss, Advisory Committee

Dr. Robert Biggin, Advisory Committee

Dr. Carmen Coleman, Advisory Committee

Dr. Marco Muñoz, Advisory Committee

Dr. Tom Shelton, Advisory Committee

Dr. Jerry Pogatshnik, Dean of Graduate School

STATEMENT OF PERMISSION TO USE

In presenting this dissertation in partial fulfillment of the requirements for a Doctorate in Leadership and Policy Studies degree at Eastern Kentucky University, I agree that the library shall make it available to borrowers under rules of the library. Brief quotations from this dissertation are allowable without special permission, provided accurate acknowledgment of the source is made.

Permission for extensive quotation from or reproduction of this dissertation may be granted by my major professor, Dr. Charles Hausman, or in his absence, by the Head of Interlibrary Services when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this dissertation for financial gain shall not be allowed without my written permission.

Signature Tarah Shaus

Date 11/10/16

SCHOOL OF ACCOUNTABILITY GRADES: THE INFLUENCE OF STUDENT AND SCHOOL CHARACTERISTICS

By

TARA GRIMES ISAACS

Master of Arts
Eastern Kentucky University
Richmond, Kentucky
2008

Master of Science
University of Tennessee
Knoxville, Tennessee
1995

Bachelor of Arts
University of Tennessee
Knoxville, Tennessee
1994

Submitted to the Faculty of the Graduate School
Eastern Kentucky University
in partial fulfillment of the requirements
for the degree of
DOCTOR OF EDUCATION
December, 2016

DEDICATION

This dissertation is dedicated to those on the front lines, who serve as a true voice for students and families who struggle to have their own voice heard.

ACKNOWLEDGEMENTS

The following dissertation would not have been possible without the help of many. Thank you to my chair, Dr. Charles Hausman, for his guidance, encouragement, patience and methodologies in teaching me how to approach the doctoral program. Thank you to Dr. Bliss for chasing me down after walking out of a statistics class to assure me that we would get through the coursework together and meeting to review work despite an emergency evacuation on campus… that was way above and beyond. Thank you to Dr. Robert Biggin for feedback, answering questions and allowing time for me to talk about the work. Thank you to Dr. Carmen Coleman and Dr. Tom Shelton, for their encouragement, mentoring, and assistance at critical points in my career. Thank you to Dr. Marco Munoz for pushing me to think more deeply about the research design and offering the right words at the right time. Thank you to Dr. Matthew B. Courtney (once my 4th grade student), Dr. Lisa Revel, Dr. Kate Grindon, Dr. Lizette Rogers, Dr. Darlene Spurlock, and Karen Branham; all who assisted with a portion of the journey. Special appreciation to my husband, Aaron, for his patience, sacrifices and presence. Many hours were given in silence sitting in the same room with me as I worked. I would also like to thank my parents, Donnie and Nancy; sister, Jamie; aunt, Rhonda, and various family members for encouraging me to finish strong. Thank you to Pastor Don for being a pastor to me even after moving to another city. Finally, I would like to thank God. Without Him, this dissertation would not have been possible.

ABSTRACT

In the United States of America, several states passed legislation that enacted a grading system by which schools are measured for their performance through a formalized ranking system, which deem schools a success or failure. These accountability systems are developed by legislators and policymakers in order to fulfill federal requirements like the No Child Left Behind Act (No Child Left Behind [NCLB], 2002) and the Every Student Succeeds Act (2015), to encourage systemic improvement across states. In reporting annual school performance for the state of Utah, for example, it is necessary to investigate to what extent student and school characteristics predict school accountability grades within their state's educational accountability system. This study utilized school-level data to assess the effect of school predictors on low SES, English Language Learners, students with a disability, and racial minorities on school accountability grades. The school characteristics that were evaluated included teachers with a graduate degree and rural locations. Prior research focused primarily on individual predictors of variance on student achievement, while this study combines all of the predictors for observation on predictors of variance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Overview

States have passed legislation to enact accountability systems by which schools are measured for their performance and ranked. Lawmakers and policymakers apply accountability system formulas in the form of grades and rankings to warrant legislation, such as No Child Left Behind Act (NCLB, 2002) and Every Student Succeeds Act (2015), in efforts for universal advancement of schools across the country.

> States must design systems of school report cards based on the fraction of students demonstrating proficiency in reading and mathematics. Under NCLB, if students do not make adequate yearly progress, schools and districts face consequences such as mandatory public school choice and the possibility of complete school restructuring, as well as the redirection of federal funds; states risk the loss of federal administrative dollars. Additionally, the classifications or grades formally assigned to schools may affect the attractiveness of the local area to potential and current residents and the perceptions of local officials by the public. (Figlio & Getzler, 2002, p. 1)

Figlio and Lucas (2000) offered evidence that housing markets are highly reactive to government-based report cards, thus inadvertently rewarding schools for focusing on accountability items resulting in less time given to school subjects not covered or providing test prep—essentially "gaming the system." In deciphering annual school performance, it is imperative for statesmen, scholars, parents, and community members to understand how states arrive at these findings, as well as understand the repercussions for specified approaches. If policymakers advocate for accountability in the form of

grades, they may not be aware that these approaches could be perpetuating additional problems that indirectly result in schools being less likely to discourage poorer students from dropping out of school (Figlio & Getzler, 2002).

Across the United States, school performance results are the report cards for predicting the success of public schools. This study investigated to what extent student demographics and school characteristics predict school accountability grades within Utah's educational accountability system.

Specifically, this study used state-level data to determine if student demographics and school characteristics are predictors of school accountability grades. The variables analyzed were low SES, minority, English Language Learners, disability, teachers with a graduate degree, and rural setting locations.

Purpose Statement

Research can isolate individual predictors for success of students, but it has not provided an approach where multiple predictors are used to determine the success of students resulting in grades assigned on a state accountability system. Additional research is needed to in order to determine the extent to which school accountability systems are influenced by student demographics and school characteristics. Additional questions arise as to whether the types of accountability models fairly assess student achievement. Using a study of accountability frameworks, student demographics, and school characteristics, the researcher sought to determine whether or not school accountability grades are prone to patterns of predictability.

Dorn and Ydesen (2014) stated that researchers look towards connections at all levels in society in order to understand root causes of changes in accountability, as well

as trace the impact and connections pointing to "questions of power, education access, educational management, and social selection" (2014, p. 2), especially as it relates to the makeup of the team tasked with designing and administering educational accountability structures. This research sought to determine if assigning grades to schools is an equitable practice. Based on prior experiences and research within coursework, the researcher expected to find patterns and predictability in assignment of grades based on a combined group of student demographics and school characteristics.

Statement of the Problem

Accountability models that assign grades to schools are used to classify overall school performance. If there is a correlation between state assessment results and student demographics and school characteristics within educational settings or specific type of an accountability model, grades could be promulgating social injustice.

Research Question

To what extent do student demographics and school characteristics predict school accountability grades? The research question examined whether student demographics and school characteristics impact grades assigned to schools. Student demographics include: low SES, English Language Learners, disability, and racial minority on school accountability grades. The school characteristics that were evaluated included teachers with a graduate degree and rural locations.

Significance for the Study

The study is significant because of the importance placed on the grades assigned to schools in this study and across the United States. Accurately understanding the impact of student demographics and school characteristics across various accountability models

is of the utmost importance due to the host of implications associated with the reported school grades. Within today's educational reform movement, significant efforts have been put into place to improve the overall quality in public education, considering prior efforts and lack thereof, incited the civil rights movement and a "growing and widespread concern about the vital importance of education to our national security" (Kress, Zechmann, & Schmitten, 2011, p. 188). In some research, the United States' educational system is no longer considered among the highest ranked countries in the world; if they are to reclaim their place among the ranks, there must be a concerted effort to accurately understand, assess, and communicate accountability models (Traylor, 2013, p. 8). Also noted by Traylor is the importance of understanding the implications of high-stakes accountability models:

> The study of the implications of different high-stakes accountability models using the same student data is imperative to current education reform because we must ensure our educational systems are identifying the schools and districts that are best producing students with the 21st century skills needed to compete globally and ensure the economic success of the United States. Further, we must be able to accurately identify those schools and districts not producing students with these skills because research strongly suggests these students will not only most likely be required to take remedial courses if they pursue post-secondary education, but they will earn significantly less income over the course of their lifetime. (Traylor, 2013, p. 8)

In a recent study, the income earning level of adults was shown to be directly connected to the educational level they attained in school. Economic variables in this

study will likely show statistically relevant connections to overall school performance. For example, Table 1.1 illustrates the discrepancy between income and education. Additionally, *The Hamilton Project* reports almost 80% of students who drop out of high school made less than $30,000 in 2010, and 80% of those students who went on to graduate from college earned around $100,000 (Greenstone, Harris, Li, Looney, & Patashnik, 2012).

Table 1.1

Income Over Course of Lifetime Based on Education

Education Level	Average Lifetime Earnings
Professional degree	$4.4 million
Doctoral degree	$3.4 million
Master's degree	$2.5 million
Bachelor's degree	$2.1 million
Associate's degree	$1.6 million
Some college	$1.5 million
High school graduate	$1.2 million
Non-high school graduate	$1 million

Source: Greenstone, M., Harris, M., Li, K., Looney, A., & Patashnik, J. (2012, September). *A dozen economnic facts about K-12 education.* Retrieved from http://www.hamiltonproject.org/assets/legacy/files/downloads_and_links/THP_12EdFacts_2.pdf

Rationale for the Study

According to Colvin and Helfand (2000), the educational system in the United States undergoes ongoing criticism for its lack of results-based reform that positively impacts student achievement. Schools, leadership, and teachers have an increasing

burden to meet weighted benchmarks in performance for accountability testing requirements. Most notably, in some instances, schools are given monetary rewards or schools are completely redesigned with staff reassigned for not meeting expected outcomes (Colvin & Helfand, 2000). In response to growing public unrest, and in part to advance student achievement mandates, public pressure has led to increased accountability for schools in the form of report cards that outline school grades through complicated accountability formulas (Ladd & Walsh, 2002).

Schools with high concentrations of Caucasian students with high socioeconomic (SES) status typically scored higher on standardized assessments than students from schools with higher volumes of English Language Learners, low SES students, and students with disability (Toutkoushian & Curtis, 2005). Grading systems that are reported to the public cause shifts in preferences for school attendance and affect geographic migration. This ultimately segregates communities, while inadvertently creating ~~more diverse~~ settings with higher concentrations of needs and hurdles to overcome (Glynn & Waldeck, 2013).

States continue to use grades to determine school performance by assigning weights to variables within accountability models that are measured, many of which cannot be controlled by the schools. Additional research into high-stakes accountability systems is necessary for state leadership, the public, and those within education to be able to correctly address challenges facing schools.

CHAPTER 2: LITERATURE REVIEW

Background

Since the end of the 1990s, various countries across the globe have increased policies to hold both schools and school districts accountable for academic performance (Dorn & Ydesen, 2014). According to Sahlberg (2010), multiple international ranking systems are increasingly widespread pertaining to the rate and rise of today's accountability as part of the new Global Education Reform Movement (GERM), which has now overshadowed the history predating this movement (Dorn & Ydesen, 2014).

Accountability grading models are being used to determine student and school success. During the introduction of state accountability in the 1990s, states that introduced consequential accountability systems early, which included both rewards and punitive actions, tended to show initial gains (Hanushek & Raymond, 2003a; 2003b). If student and school variables can predict state accountability results, grades could be contributing to a social injustice. There is little debate as to the importance of a strong education and its impact on student success.

With state testing accountability performance and reform among the forefront of issues in education today, many schools are judged by an assigned grade based on an accountability formula. Parents, community members, and media use the grades to define the academic performance of the school in which to determine the best schools that will provide students a greater chance at success. "Without a doubt, the achievement of our students has direct ramifications for the future well-being of our society" (Hanushek, 2004, p. 323). Questions remain as to whether grading accountability systems are set up to further separate those that are advantaged from those who are disadvantaged.

Federal Accountability

In 1965, the Elementary and Secondary Education Act (ESEA) was passed to "ensure equal educational opportunity for all children" and "to close the achievement gap between poor and affluent children" (Elementary and Secondary Education Act [ESEA], 1967).

Current educational trends in accountability have been the focus instead of the previous debate over testing and accountability origins dating as far back as the 1830s and 1840s (Reese, 2013). Beginning in 2001 under the No Child Left Behind Act (NCLB, 2002), President George W. Bush signed a law requiring all states to provide an annual measure of learning through statewide assessments in grades 3-8 and identify disciplinary measures for those not making academic gains in designated periods of time (Robelen, 2002; Stecher & Hamilton, 2002). The No Child Left Behind Act of 2001 (NCLB, 2002) requires states to monitor student and school performance based on "adequate yearly progress" (AYP), which is essentially a count of the number of students meeting a specified target.

> According to Goldschmidt and Choi (2007), NCLB also requires that
>
> 100% of students must demonstrate proficiency in reading and mathematics by 2013-2014 for those schools receiving Title I funding. Furthermore, schools must demonstrate adequate yearly progress towards the 100% proficiency target. A school that does not meet the annual target (set by each state) faces increasingly severe sanctions based on the number of contiguous years that he school misses its target. NCLB presumes that monitoring the percentage of students who are

> proficient in reading and mathematics is sufficient to identify schools that are doing a good job and schools that need improvement. (pp. 2-3)

Under the No Child Left Behind Act (NCLB, 2002), schools that fail to meet adequate yearly performance (AYP) targets for two consecutive years are labeled as "failing." This general term produces a negative association for schools that are most often in low-income communities, which are the same geographic areas legislation intended to assist through the NCLB Act of 2001 (2002). "Failing" labels result in significant decreases in home and property values due to overarching "perceptions of poor school quality or social stigma surrounding a 'failing' designation" (Bogin & Nguyen-Hoang, 2014).

"To respond to those types of pressures, states will have to devise criteria for identifying schools that they deem as underperforming and in need of improvement," which will primarily focus on student performance (Toutkoushian & Curtis, 2005). Under this criteria, policymakers are comparing schools based on student performance that do not take into consideration factors that are beyond the control of schools (Toutkoushian & Curtis, 2005). Due to NCLB federal legislation, all states are required to report an Annual Measurable Objective (AMO). For example, in Utah,

> Based on the percent of student achieving proficiency on the states [English Language Arts (ELA)] and mathematics of the [Student Assessment of Growth and Excellence (SAGE)], [Annual Measurable Objectives (AMO)] targets are set for each school and subgroup in annual equal increments toward a goal of reducing by half the percentage of students in the all students group and in each

subgroup who are not proficient within six years. AMOs are reported for the following groups:

- All Students
- Economically disadvantaged
- English learner
- Hispanic/Latino
- Students with disabilities
- White

(Utah State Office of Ed, 2015, pp. 24-25).

In 2015, the Every Student Succeeds Act was signed into law, providing states with an about-face from its predecessor, the No Child Left Behind Act. In this legislation, the United States Department of Education would play a more limited role in accountability, allowing states the flexibility to choose their own goals as long as they address "proficiency on tests, English-language proficiency, and graduation rates," along with "an expectation that all groups that are furthest behind close gaps in achievement and graduation rates" (Education Week, 2016).

At the elementary school and middle school levels, states are required to incorporate a minimum of four indicators, which include the three mandated indicators: a) proficiency on state tests; b) English-language proficiency; and c) graduation rates—plus another academic variable "that can be broken down by subgroup, which could be growth on state tests" (Education Week, 2016). High schools are held to the same standards as the elementary and middle schools, except that they must also include graduation rates. Ultimately, individual states will determine how much each indicator

will count, although the graduation rates must be weighted heavier (Education Week, 2016).

Under ESSA, states must pinpoint and intervene for the bottom five percent of student performers, high schools where graduation rates are 67% or lower, and schools where subgroups of student population are struggling academically (Education Week, 2016). For schools that require interventions, the following will be applied for the bottom five percent of schools and for high schools with high dropout rates:

- Districts will work with teachers and school staff to create an evidence-based plan.
- States will monitor the turnaround effort.
- If schools continuously fail, after no more than four years the state will be required to step in with its own plan. A state could take over the school if deemed necessary, or fire the principal, or turn the school into a charter school.
- Districts could also allow for public school choice for seriously low-performing schools, but they must give priority to the students who are in most need (Education Week, 2016).

For schools where subgroups of students are struggling:

- Schools must create an evidence-based plan to help the particular group of students who are falling behind, such as minority students or those students with disabilities.
- Districts must monitor these plans. If the school continues to fall short, the district would step in, though there's no specified timeline.

- Importantly, there is also a provision calling for states and districts to create a "comprehensive improvement plan" in schools where subgroups are chronically underperforming, despite local intervention.
- The School Improvement Grant (SIG) program is consolidated into the bigger Title I pot, which helps districts educate students in poverty. States could set aside up to 7%of all their Title I funds for school improvement, up from 4%in current law (Education Week, 2016).

Tests are required by states in reading and math for grades 3-8 and once in high school for a disaggregation of data for the whole school along with all subgroups including: English-learners, disability, racial minorities and economically disadvantaged as well as maintaining a 95%participation rate of students (Education Week, 2016).

As with NCLB, new ESSA legislation continues to task, monitor, and punish schools, school leaders, and districts that are not able to meet legislation standards and ultimately provides the state level with leverage to fire school leaders, take over the school if needed, or turn the school into a charter school when turnaround efforts fail.

Theoretical and Conceptual Frameworks

Matthew Effect

In educational research, one phenomenon or circumstance studied is known as the "Matthew Effect," which refers to the positive connection between the initial reasoning and thinking of a person and the continual building on the original starting point over time, leading to more wide-ranging amassed information (Stanovich, 1986; Walberg & Tsai, 1983). This phrase originated from a Biblical assertion: "To all those who have, more will be given, and they will have an abundance; but from those who have nothing

even what they have will be taken away" (Matthew 13:12). The use of the phrase in current times is credited to a sociologist named Robert Merton (1968), who made the point that well-known scientists were given more acknowledgement than those scientists who were lesser-known, even though both might have done similar work in their respective fields of study. Merton continued by noting a "cumulative advantage" where "the rich get richer" (p. 62). Essentially, Merton drew parallels with the "Matthew Effect" phenomenon and the advantages or disadvantages that students bring into their education as a predisposition of how far students could achieve in learning and cognitive growth.

Other authors have discovered large-scale patterns among literacy and math development, where each student's abilities and differences increase as their initial standing of where they rank among those being compared to over time (Bast & Reitsma, 1998; Juel, 1988; Seltzer, Frank, & Bryk, 1994; Stanovich, 1986). Even with consistent and continual growth, students who begin their academic career at a disadvantage will remain at a disadvantage without significant outside intervention. Unless disadvantaged students make more growth than those without advantages, it is unlikely to see a complete gap closure between groups. Additional findings from various authors indicated "one-sided" Matthew effects, where those students who are low achieving continue to fall behind during the first few elementary years, while gaps in learning remain the same with high-achieving and average students (Morgan, Farkas, & Hibel, 2008).In other words, there will always be an achievement gap that is dependent on where students begin their education.

Bandura's Theory of Self-Efficacy

Student body characteristics influence school-level achievement more strongly by altering faculties' beliefs in their collective efficacy than through direct effects on school achievement (Bandura, 1993, p. 117). Figure 2.1 indicates a "path analysis showing the role of perceived collective efficacy in the casual structure of school-level achievements in reading and mathematics" (Bandura, 1993, p. 143).

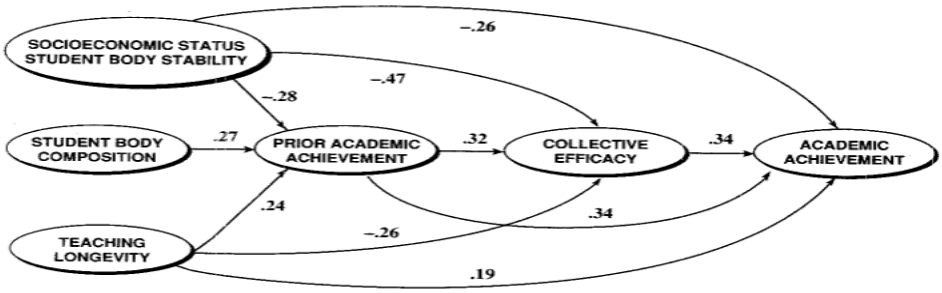


Figure 2.1. Bandura's Theory of Self-Efficacy

Source: Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, *28*(2), 117-148.

Bandura (1993) noted the following:

> Adverse student body characteristics influence schools' academic attainments more strongly by altering faculties' beliefs about their collective efficacy to motivate and educate their students than through direct effects on school achievement. Indeed, with staffs who firmly believe that, by their determined efforts, students are motivatable [*sic*] and teachable whatever their background, schools heavily populated with minority students of low socioeconomic status achieve at the highest percentile ranks based on national norms of language and mathematical competencies. (p. 143)

Fan-Close Theory

In comparison to the Matthew Effect, the beginning academic performance of lower-achieving students may appear to show more progress than high performing and average students (Ready, 2012). In reviewing an extensive assortment of assessments, experts have described a closer spread or "fan-close" connection at the elementary level (Ready, 2012, p. 94), which may be related negatively to academic gains. Studies on the Tennessee Value-Added Asessment System also determined a connection that on average, initial higher-achieving students demonstrated lower gains over time (Wright, Horn, & Sanders, 1997). Similarly, studies using assessment data from North Carolina showed the same results (Rothstein, 2008a; 2008b). Therefore,lower-performing students may have the appearance of making more growth as compared to higher-performing students because of more learning ground to make up after beginning school. However, this may not be a fair assessment of growth between lower-performing and higher-performing students.

Measuring the connections between academic performance and student growth may correctly depict logical, evolving progressions (Ready, 2012). At any given point in measuring progress in abilities and learning, a connection—positive or negative – may be present between where students begin learning and how fast they are able to learn (Ready, 2012). On the other hand, these findings could appear to show relevance, but may be false in that they merely seize distinctive psychological variables and a maneuvered methodological properties from certain assessments (Ready, 2012). Analytical methods can address profoundly diverse questions, resulting in different results depending on how experiments are designed (Ready, 2012).

Proponents who argue for the importance that relationships play in impacting student performance point to both individual and social variables that may determine whether or not students grow in learning (Ready, 2012). For example, students who excel at reading may receive more positive feedback and praise than those who do not, which in turn creates more inclination to learning through increased incentives and self-worth (Ready, 2012). When viewing this through the lens of the "rich get richer" idiom, where beginning abilities are clearly related to acquisition of language, students with more vocabulary in their background tend to read more, acquire more vocabulary, and, as a result, read better, while lower performing readers read slower and with less pleasure, so they tend to take fewer opportunities to read independently, thereby resulting in a slower pace of growth (Bast & Reitsma, 1998; Stanovich, 1986).

In reviewing students' academic backgrounds, in relation to family and community; socioeconomic variables connect to both by perpetuating enjoyment and affirmative links to learning for high-performing students as they move through school due to positive mental experiences and continual growth (Ready, 2012). These essential effects may be visible in circumstances where ability grouping is practiced, where students are separated into advanced groups and put into configurations where they perceive, due to their ability, to learn more (Pallas, Entwisle, Alexander, & Stluka, 1994). Consequently, instruction taught in similar groupings may not reveal constructive relationships between the school in which a student begins their learning and end results for growth (Ready, 2012). Matthew Effects may also be prevalent due to disparities in resources for personnel and materials, which have tended to be superior in settings that house larger numbers of high-performing students (Ladd & Walsh, 2002).

Other practices and procedures may suggest an adverse relationship between where students begin their learning and how much growth they are able to make (Ready, 2012). For example, students who do not speak English as their first language would theoretically gain English abilities at a quicker pace than non-English Language Learners (ELL) peers, but might test lower on a separate assessment (Ready, 2012). A separate justification from advocates for gifted and talented programs is that instruction and curriculum may not be rigorous enough for high-performing student needs (Ready, 2012). Academics have recognized undesirable connections with where students begin their learning and growth in academics with the Tennessee Value-Added Assessment System in that there is a "lack of opportunity of high-scoring students to proceed at their own pace, lack of challenging materials, lack of accelerated course offerings, and concentration of instruction on the average or below-average student" (Wright, Horn, & Sanders, 1997, p. 66).

Connections between where a student begins in learning and where they reach with growth may also stem from theory and techniques known as *psychometrics*, where skills, knowledge, approaches, aptitudes, and personality traits are variables that are used for assessments to measure growth (Koedel & Betts, 2008; Raudenbush, 2004; Smith & Yen, 2006). Three significant areas that are related to the regression toward the mean are important to note when studying growth (Ready, 2012). *Regression toward the mean* is a term that indicates that a measurement is extreme on its first test and more likely to be close to the mean on the second test—or, if extreme on the second measure, it is accepted that it was probably closer to the average on the first measure. The measure can be comparatively attributed to chance (Lane, n.d.; Regression, n.d.).

The first concern consists on whether or not state tests have the capacity to accurately assess growth among high-achieving students. The bank of questions used for assessments may not introduce sufficient quantities of rigorous questions over time, thus producing mean regression, and a false-negative correlation between beginning knowledge and growth. In these conditions, low-achieving students may appear to gain more academic skills on assessments over time. This seems quite conceivable, as many state assessments focus on measuring border line proficiency levels rather than growth among high-achieving students (Koedel & Betts, 2008).

A second justification for connections between student achievement and growth has to do with errors in measurement characteristic in all standardized tests, in that students who perform toward the ends of tests may have done so by accident, indicating that scores of these students are likely to reach the average if given time, thus creating a negative relationship between beginning knowledge and growth (Ready, 2012). To account for this problem, researchers use a two-stage least squares approach, which allows for averages of students' test scores in other subjects to be used as other relevant variables in value-added models (Booker & Isenberg, 2008). Studies of data from state assessments indicate that taking this error into consideration with correction radically changes school rankings, due to the randomness of assignment of students by achievement areas (Ladd & Walsh, 2002). Further examinations and studies from diverse settings are in agreement, as well as note that the error adjustment by itself cannot fully correct the relationships between beginning learning and growth made over time in academics (Phillips, 2000; Rothstein, 2008b).

In reviewing three dissertation studies, Alcorta (2011), Traylor (2013), and Phillips (2015), assessed the influence or relationship between student demographics or community traits and even touched on the ways in which data is run and/or how accountability models and formulas factor into the discussion. Additionally, the connection between growth, achievement, and gap scores were essentially found to be a repeated data point based on the specific demographic in question. Of great interest were the four specific types of accountability models that were explained in Traylor's (2013) dissertation, each of which were different. Traylor noted the following accountability models: growth, status, value-added, and improvement. Each of them were based on a different measurement of student growth. Growth indicated whether or not student's performance changed. Status observed an average of how students were performing. Value-added focused on whether or not students' changes in performance was enough to meet the growth expectation and whether or not students met or missed the target. Improvement considered on average how were students doing during the year compared to students in the same grade the previous year.

Interestingly, the grading/ranking of student and school performance would report different data depending on the perspective of the accountability model. Phillips (2000) also determined different outcomes for grades/ranking based on score transformation within measures of varying demographics. Finally, Alcorta (2011) replicated the closest study and experiment to what the researcher proposed in reviewing the impact of community traits, school attributes, and student demographics as it relates to student achievement. Variables differed as compared to the original plan of the researcher, causing minor tweaks in eliminating the types of accountability models and adding

variables for school community factors. Traylor (2013) did not list any theoretical frameworks, although there were some available to connect with accountability model topics that the researcher recently found.

When viewing accountability models and the grades assigned to schools as a result of the model, it is essential to understand formulas that are used to accumulate total points. Not only is the accountability model development important to understand for context, but equally important are the weights and items taken into consideration for measurement. In some situations, as in the Utah data, students are given more of an advantage if they graduate and earn fewer points based on ACT performance, which has a direct correlation to whether or not a student is able to be accepted into college. It is essential to avoid penalizing or rewarding unfairly any set of students due to student demographics and/or school characteristics.

Alcorta (2011) provided a conceptual framework, shown in Figure 2.2 that outlined student demographics, school attributes, community traits, and instructional expenditures all relevant to student achievement.

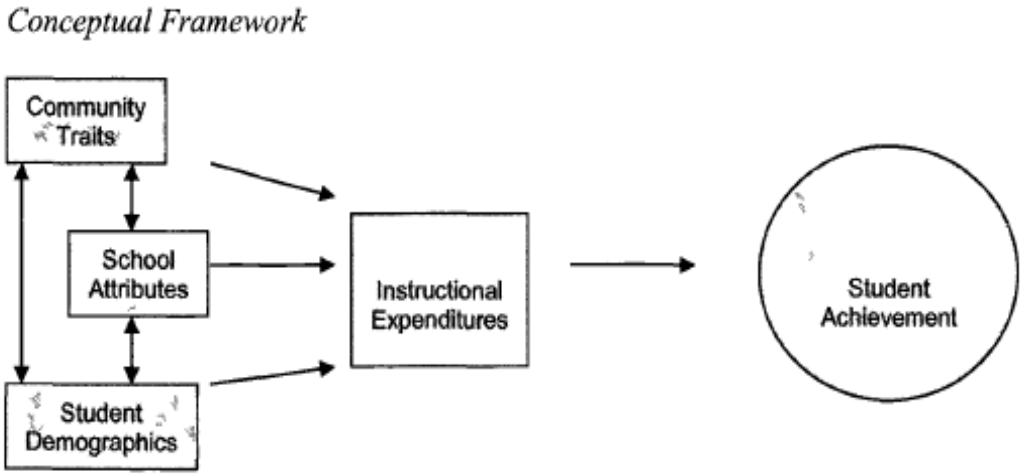


Figure 2.2. Alcorta's Conceptual Framework

Source: Alcorta, J., Jr. (2011). The impact of community traits, school attributes and student demographics on student achievement. *Tarleton State University*, 1-150.

Phillips (2015) explored methods of transformational measurement and the influence that particular demographics had on final measurement outcomes. Overall, each of the three dissertations provided context through connections to similar subjects of the researcher, provided a framework for replication, and provided information about student demographics and school community factors that added to the body of current literature. This conceptual framework example sharpens the focus on accountability by reviewing not only at performance, but also on the physical settings in which student demographics and school characteristics are being evaluated by to determine if there are implications for student results.

Student Demographics as Influences

Low Income Students

Researchers note that society accepts that SES directly impacts the performance of students within schools (Toutkoushian & Curtis, 2005) and this is supported by Bracey's (2004) comment that "poverty is not an excuse, it is a condition. Like gravity, it affects everything" (p. 636). Wake (2001) asserted that people can predict the performance of the school by the cars in the parking lot. In other words, the more expensive the vehicles in the school lot, the better the school's performance and reputation with district prestige. In a study of 18 separate school outcome measures in Illinois, Fowler and Walberg (1991) found that "District socioeconomic status and the percentage of students from low-income families in the school were the most influential and consistent factors related to schooling outcomes" (p. 189).

In a national climate of high-stakes testing, there is especially great pressure on schools with high concentrations of low performing students to demonstrate academic growth (Moon, Callahan, & Tomlinson, 2003):

> Skeptics have begun to wonder if the effort to raise standards for all students through high-stakes testing initiatives has too steep a price, including a narrowing of the curriculum and a de-emphasis on curricular depart, an abandonment of constructivist-type activities that give meaning to learning, and a curtailment of extracurricular activities. (Moon, Callahan, & Tomlinson, 2003, p. 1)

Despite teacher credentials and content knowledge, teachers in high-poverty schools tend to have to work with outdated materials, technology, and supplies, along

with offering significantly fewer college preparatory or advanced placement courses as compared to more advantaged populations (Freel, 1998).

In addition, low teacher expectations also may play a part in affecting teacher instructional strategies in working with economically disadvantaged children (Moon, Callahan, & Tomlinson, 2003). Researchers have further found that "teachers tend to have lower expectations for students from improvished [*sic*] backgrounds and they often formulate these expectations before they have significant interaction with students" (Moon, Callahan, & Tomlinson, 2003, p. 2). Teachers often provide basic or watered-down instruction (Ornstein & Levine, 1989), "reinforcing the drill-and-practice of basic skills while ignoring higher-order thinking skills that enable complex and meaningful learning to occur" (Moon, Callahan, & Tomlinson, 2003, p. 2). This teaching behavior essentially devalues academic chances for students (Moon, Callahan, & Tomlinson, 2003).

In their study of 150 public schools in Pennsylvania, Summers and Wolfe (1976) found a notable link between SES factors and student performance. Michelson (1972) discovered that the percentage of elementary students who were eligible for free or reduced-price meals and average family income were related to reading performance. Similarly, Toutkoushian and Curtis (2005) indicated the following:

> The literature on student success at the K-12 and postsecondary levels has shown convincingly that socioeconomic factors are correlated highly with student, and hence school outcomes. That relationship could be caused by differences in underlying family characteristics that affect student performance. Although the statement that SES has a causal effect on performance might not be precise, that is

> the convention adopted by most analysts, perhaps as a matter of convenience (p. 261).

Toutkoushian and Curtis (2005) pointed out that "policymakers should be concerned that comparisons of schools that are based on outcome measures do not consider the existence of factors that are correlated highly with these outcomes, many of which are beyond the control of the school" (p. 260). Other researchers determined that there are other factors connected to SES, such as parent education and income, which are strongly connected to student performance at each educational level (Stinebrickner & Stinebrickner, 2003; Summers & Wolfe, 1976; Sutton & Soderstrom, 1999). "Higher levels of student performance may not be caused by a community's SES per se, but, rather, by factors that are correlated with SES, such as parents' emphasis on education" (Toutkoushian & Curtis, 2005, p. 260).

As part of a study on student performance on standardized assessments in the Illinois public schools, Sutton and Soderstrom (1999) determined a strong association between student performance and the income level and ethnic diversity of a community. Jaggia and Tuerck (2000) indicated that the SES of communities was a predictor of test scores of students in Massachusetts. The authors specified that "what matters most for the current performance of a school district is its past performance and the socioeconomic character of the district. We cannot make schools perform better just by spending more money on them" (p. 4).

In New Hampshire, Hall (1998) applied descriptive statistics to scores for third grade students and found that performance varied by family income and educational level of parents. Toutkoushian and Curtis (2005) asserted that although correlations can be

found in statistics between income and education levels' impact on student performance, "it is not universally accepted that schools in low SES regions cannot perform at high levels" (p. 261).

The Education Trust (Jerald, 2001) delivered evidence to dispute any positive association between academic performance and low performance for students living in low-SES areas. According to Toukoushian and Curtis (2005), researchers have continued to be baffled as to why researchers have not been able to determine a dependable set of factors from empirical studies that prove that student performance on assessments can be controlled by the school and community. "Factors such as the wealth and educational status of residents and the state of the local economy are for all intents and purposes beyond the control of schools" (Toutkoushian & Curtis, 2005, p. 261).

On the other hand, Jaggia and Tuerck (2000) found instances where schools had low test scores, but outperformed expectations given their SES, and cautioned that failure to consider the SES in schools could lead to situations where "schools with low ratings but good administrators and teachers will be falsely perceived as doing a poor job of teaching their students" (p. 1). Furthermore, Jaggia and Tuerck (2000) posited that much attention has been given to high school student test scores in Massachusetts, but was not necessarily helpful to policymakers because of "socioeconomic factors over which policymakers and educators can exert little control but that nevertheless are highly important in determining how individual districts perform" (p. 37).

Racial Minority Students

Although researchers have guessed that high-stakes testing had more of a negative impact on minority students than other students, very little attention has been given to

investigating the impact of these negative effects on students, especially those prone to dropping out of school or at risk of failure or those who live in urban settings (National Commission on Testing and Public Policy, 1990). Oakes (1990) noted, "One impact of mandated standardized testing on minority students cited by a growing number of researchers is its role in the denial of opportunities to minorities" (p. 172). O'Connor further asserts, "Much of the literature on testing and minority students deals with decisions based on test performance, but some research has focused on the probably sources of differential performance such as language, social, and cultural factors" (1989, p. 172).

Regardless of the amount of research given to researching the disparities between minority and non-minority students, little headway has been made by a variety of stakeholders (Kulm, 2007).Researchers have presumed the existence of evidence that achievement gaps exist between white and minority students even prior to entering kindergarten (Chapin, 2006). Irrespective of what causes the achievement gap, the focus must center on providing research-based interventions to correct the problem (Williams, 2011). Minority status alone does not necessarily indicate low achievement. However, minority status, compounded by other factors such as low SES, are indicative of lower achievement (Fowler & Walberg, 1991).

Despite conventional wisdom that school inputs make little difference in student learning, a growing body of research suggests that schools can make a difference, and a substantial portion of that difference is attributable to teachers. Recent studies of the effectiveness of teachers at the classroom level using the Tennessee Value-Added Assessment System and a similar data base in Dallas, Texas, have found that differential

teacher effectiveness is a strong determinant of differences in student learning—far outweighing the effects of differences in class size and heterogeneity (Sanders & Rivers, 1996; Wright, Horn, & Sanders, 1997; Jordan, Mendro, & Weersinghe, 1997).

In 1954, school segregation came to the frontline in educational policy when the U.S. Supreme Court determined that by law segregation based on race was unconstitutional (Brown v. Board, 1954). Multiple court orders followed in attempts to integrate large urban school districts through voluntary choices and involuntary means, such as student reassignment or busing plans (Armour, 2003; Frankenberg, Lee, & Orfield, 2003). Whether *de jure* or *de facto*, achievement gaps between white and minority groups remain well-documented on almost every measure of student achievement (Olszewski-Kubilius, Lee, Ngoi, & Ngoi, 2004).

> Conceivably, racial/ethnic residential segregation could be a reflection of the large socioeconomic status (SES) gaps that exist between members of US minority groups and Whites at the individual level. However, this does not appear to be the case. The available evidence indicates that segregation by race/ethnicity is stronger than segregation by income, that is, race and ethnicity sort individuals of comparable SES into vastly different neighborhood environments. (Acevedo-Garcia, Lochner, Osypuk, & Subramanian, 2003, p. 215)

English Language Learners

With the arrival of the No Child Left Behind Act (NCLB, 2002), United States federal law requires the inclusion of English language learners (ELL) and students with disabilities in large-scale assessments in order to provide accountability to schools for groups that are specified under the regulation (Pitoniak & Royer, 2001). High-stakes

assessments given to a large volume of students negatively impact both ELL and students with disabilities due to the difficulty in reading assessment items with limited English proficiency or language skills (Abedi, 2004; Johnson & Monroe, 2004). In turn, what policymakers do with the testing results is of equal importance as it relates to disciplinary measures, which directly affect students' lives and future opportunities in education (Solórzano, 2008). "Achievement tests were not designed with ELLs in mind. As a result, the validity of inferences from these tests can compromise the educational decisions that educators make based on test results" (Solórzano, 2008, p. 282). These practices cause problems in correctly assessing student progress for growth, diagnostic, or program evaluation (Solórzano, 2008).

Goldschmidt and Choi (2007) assert the following:

> Simply monitoring the percentage of students in a school who score at or above the proficient level in comparison with an annual target percentage places too much emphasis on student enrollment characteristics (a school that routinely receives a large influx of limited English proficient students each year will be at a disadvantage in comparison with a school that receives very few). (p. 3)

Utah schools with ELL populations are possibly placing too much emphasis on the test scores of students with relatively new exposure to language acquisition, which calls into question equity.

> Fairness and bias issues are paramount because ELLs may be stuck in a remedial English oral language proficiency curriculum with little opportunity to learn the content and skills necessary to do well on high stakes tests or to have available accommodations that take their language proficiency levels into consideration

> during the tests. Taken together, these issues are compounded in such a way as to create a systemwide [*sic*] barrier to learning and educational progress and opportunity that threatens to broaden the gap between ELLs and the rest of the student population. This in turn, will result in promoting a generation of youth who are undereducated, tracked into lower paying jobs, and susceptible to all the negative consequences of being marginalized in today's society. (Solórzano, 2008, p. 261)

Students with Disabilities

One of the major contentions with large-scale assessments for both the ELL and students with disabilities populations is the negative impact associated with reading difficulty and math assessment items due to students having lower English proficiency or language skills (Abedi, 2004; Johnson & Monroe, 2004). In considering the increasing frequency of high-stakes testing that directly impact students with disabilities, some researchers have indicated concerns about unintended consequences, such as increases in dropout rates (Thurlow & Johnson, 2000). Some consequences include: a) an increase of referrals for special education services; b) teachers lowering expectations of students; c) limiting curriculum and instruction to only areas tested; d) overuse of test preparation materials without differentiating instruction; e) limiting students' involvement with extracurricular activities due to extra time working on areas of need; and f) the use of test scores to determine if a student will graduate with or without a standard diploma (Education Commission of the States, 1998; Lane, Parke, & Stone, 1998; Langenfeld, Thurlow, & Scott, 1997; Nelson, 1999).

> Many positive and negative consequences of high-stakes testing for students with disabilities are alleged. Yet, there is little evidence on actual consequences. Both anecdotal and empirical evidence were reviewed with regard to increased participation in assessment, raised expectations, provision of appropriate assessment accommodations, alignment of individualized education programs (IEPs) to standards and assessments, improved access to general education, improved instruction, changes in promotion and grade advancement decisions, graduation and diploma options, test stress, and improved educational outcomes. Data needed to make judgments about intended and unintended consequences of high-stakes testing are also analyzed. (Ysseldyke, et al., 2004, p. 75)

Testing is also used for determining promotion to the next grade level or retention. Some researchers have pointed out the practice of retaining students with disabilities or those who have not been able to meet grade level benchmarks as an inequitable effort to inflate test scores (Langenfeld, Thurlow, & Scott, 1997; Zlatos, 1994). If students with disabilities are asked to be a part of their state high-stakes testing process, then conversations around unapproved accommodations, along with their roles on assessments, must be further discussed and considered. Accessibility for students with disabilities must be available in order for these students to have an equitable chance at showing their learning, and working with the team that closely supports students is key to maintain high expectations for learning and growth (Thurlow & Johnson, 2000).

Other ways to ensure that students are given fair criteria for promotion decisions is to make sure that there are clear processes in place to: a) identify struggling and/or at-risk students (students with disabilities); b) allow for multiple points of accessing

curriculum and demonstration of learning (differentiation of content, process and product); and c) take into consideration all of the information that is relevant to the student's learning to avoid an over identification of students that may struggle during different times of their school year (Thurlow & Johnson, 2000).

> The consequences of high-stakes testing for student with disabilities, particularly of tests used to determine graduation status or type of diploma, last well beyond the time a student is in school participation in postsecondary education programs, employment and future earnings, civic participation, and the individual's overall social and emotional well-being are affected by the credentials they receive in high school and carry forward into adulthood. (Thurlow & Johnson, 2000, p. 312)

School Characteristics as Influences

Teachers with a Graduate Degree

For many years, educators and policymakers have discussed which school variables have the greatest impact on student achievement (Darling-Hammond, 2000). Darling-Hammond (2000) found that within both qualitative and quantitative research studies, policymaking in the form of continual investment in teacher quality may be correlated to advances in student achievement. "While some evidence suggests that better qualified teachers may make a difference for student learning at the classroom school and district levels, there has been little inquiry into the effects on achievement that may be associated with large-scale policies and institutional practices that affect the overall level of teachers' knowledge and skills in a state or region" (Darling-Hammond, 2000, p. 2).

> An effective teacher receiving student from a relatively ineffective teacher can facilitate excellent academic gain for his/her students during the school year. Yet

> these analyses suggest that the residual effects of relatively ineffective teachers from prior years can be measured in subsequent student achievement scores. (Sanders & Rivers, 1996, p. 4)

In Sanders & Rivers's (1996) findings, they reported a 50-pointpercentile difference in a three-year sequence. The researchers noted that the impact of teachers on student achievement were "additive and cumulative," and as teacher effectiveness improved, lower performing students were among the first to improve and that students with varying ethnicities responded to instruction similarly (Sanders & Rivers, 1996, p. 2).

One staggering piece of data is that America's most qualified teachers are least likely to be found within the highest-poverty schools (Hundley, 2013).

> Teacher choice of school can also complicate the estimation of teacher effects. . . . Experienced teachers frequently have an option to move across districts and to choose the school within the district in which they are teaching, and they tend to take advantage of this. (Greenberg & McCall, 1974; Murnane, 1981)

Hanushek, Kain, and Rivkin (2004) "further show that teachers switching schools or districts tend to move systematically to places where student achievement is higher" (p. 9). "This movement suggests the possibility of a simultaneous equations bias—that higher student achievement causes more experienced teachers or a least that causation runs both ways" (Hanushek & Rivkin, 2006, p. 9).

Rural Schools

In reviewing factors that influence school accountability grading, school location comes into question. For many years, there has been a growing trend of family farms closing or communities across the United States dwindling in population within rural

community settings. Large-scale franchises have replaced family farms in dominating the reins of America's food supply.

According to research from Carr and Kefalas (2009):

> A closer and more clear-eyed examination reveals that our country is in the throes of a most painful and unpredictable transition. In what has become an all-too-familiar story, rural states such as North Dakota and West Virginia share an unsettling problem: too many people in their twenties and thirties are leaving. Rural counties in Kansas and Georgia report the highest rates of population loss nationally, and this hemorrhaging of people, specifically the younger generation, is the hollowing out many of the nation's small towns and rural communities. (p. 1)

Farm and industry closures across America are becoming more prevalent, causing increased job losses associated with various industries (Sherman. & Sage, 2011). This weakening across rural communities, chips away at the usual steady revenue transferred into the local economy. As a result, heads of families who once were able to follow in their family's footsteps, find themselves unable to find work or make enough money to sustain a reasonable living. This drain on the economy, business productivity, resources, and human capital has created a gap in the community offering basic services and amenities. In general, the reduction of population weakens education levels of locals, suppresses growth of remaining youth by the lack of investment from the schools, and contributes to a higher number of residents living in poverty.

Within once-thriving communities that were wiped out by the closing of industries, such as in Golden Valley, California, residents of the former logging

community differ in standpoints regarding the role in education (Sherman & Sage, 2011). Some populations identify with the necessity of schools, while others see education as a method to undermine students. Often in rural settings, schools are the core of community, but they are also recognized for the separating of classes and as a source of aggressive pushes for high achieving students to move out of the community in pursuit of higher education in order to obtain better paying jobs (Sherman & Sage, 2011).

The probability of outmigration escalates with the educational level of young people, along with the draw of environmental features such as coastal lines or even tourist attractions (Kodrzycki, 2001). "State economic and quality-of-life conditions also influence migration" (Kodrzycki, 2001, p. 30).

Children of poverty are often vulnerable, as they are entirely reliant on their family for provision and survival (Egebeen & Licter, 1991). These findings would indicate that the family structures and segregation of poverty to isolated areas within rural communities have a substantial impact on children's futures within each generation (McLanahan, 1988). An overwhelming fact among rural education researchers is the "disproportionate" number of students who drop out of high school (Licter, Cornwell, & Egebeen, 1993, p. 54). Researchers have found that contact to poverty over time has a profound effect on prediction of both school and economic success as an adult. Upon closer investigation for trends, it is shown that many of these students' attitudes are shaped by family belief systems, while others are due to coming from homes where there may only be one parent.

Another principle of research studying the migration of people suggests that there is an intentional separation of people with varying economic backgrounds (Kodrzycki,

2001) with regards to areas of residence. Low income students are most often underrepresented within the top tiers of the highest performing students during high school years, which in turn leads to low enrollment in colleges (Gerald & Haycock, 2006). Within the context of migration, communities with residents left behind tend to be more isolated economically, compounding the difficulties facing rural communities (Kodrzycki, 2001). While "home ties and intervening life choices" appear to have a substantial impact on whether or not residents return to rural areas, most are lost to urban areas, but are able to bring in some urban residents to rural communities (Gibbs, 1995, p. 35).

Accountability Models

"An accountability model is a systematic method of summarizing school performance" (Goldschmidt et al., 2005, p. 17). One accountability model becomes the foundation upon which school performance is created (Goldschmidt et al., 2005). Differing purposes of accountability models often depend on who uses the results. Parents are interested in information for the purpose of enrolling their children in "good" schools. The general public wants to know how well their local schools are doing. Education policymakers use accountability results to enforce state or federal achievement goals and often to monitor school performance in order to levy sanctions or provide rewards. Whatever the use, all audiences share the common assumption that accountability results are accurate and that valid inferences and good decisions can be made based on those results. Importantly, all accountability models will likely result in some intended consequences—for example, higher test scores (Goldschmidt & Choi, 2007, p. 2).

As part of improving student achievement across school districts and states, increased accountability for schools has added public pressure in the form of school, district, and state report cards, along with complicated accountability formulas (Ladd & Walsh, 2002). Goldschmidt and Choi (2007) indicated that "policymakers must consider the purpose of an accountability model. Without knowing what policy intends to accomplish, policymakers or educators cannot choose from a myriad of model options or make valid inferences from model results" (p. 2).

Regardless of the accountability model used, no model exists that can guarantee higher achievement, even with added benefits like rewards or punitive measures and sanctions. What an *effective* accountability model can provide is "improved learning, quality decision making, and confidence in the entire accountability system" (Goldschmidt & Choi, 2007, p. 2). Figure 2.3 illustrates "the overall cumulative pattern of accountability across all states" (Hanushek & Raymond, 2005, p. 307). Additionally,

> The data are broken up into states that attach consequences to their systems and states that simply report on school achievement. To understand the estimation strategy better, the set of NAEP testing dates for eighth grade math and reading performance is superimposed on the pattern of accountability. The phased introduction across time and across the different testing periods permits disentangling the impact of accountability. (Hanushek & Raymond, 2005, p. 307)

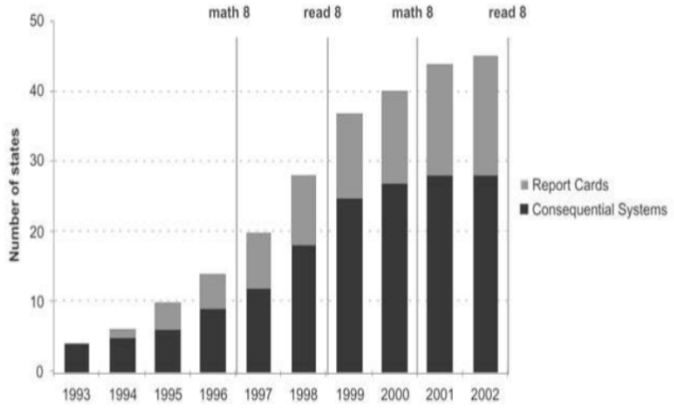


Figure 2.3. State Accountability Over Time (with NAEP Testing Dates)

Source: Hanushek, E. A., & Raymond, M. E. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, *24*(2), 297-327.

Recently, NCLB was the common element of accountability models among states. While NCLB required that states meet a target of 100% student proficiency in mathematics and reading, many details were left to each state. Operationally, this meant that while NCLB was the basis for all state accountability models, states varied in their actual design and use of this model. For example, some states had simple linear trends, while others had stair-step patterns toward 100% student proficiency.

Although a state's AYP model is based on progress toward 100%proficiency, it is not a growth model because performance is evaluated yearly based on that year's performance (Goldschmidt et al., 2005, p. 11). School accountability based on meeting AMOs defined by the status model used in AYP measures for NCLB may not correctly classify school performance. This occurs for several reasons. One reason is that schools

with more subgroups represented are more likely to miss meeting AYP due to the greater numbers of AMOs they need to meet (Novak & Fuller, 2003, p. 11).

A second reason for misclassification is that classifications based on a cut score capture only a small proportion of students' performance especially when scores are close to the cut score (Thum, 2003). For example, a student whose scores are one point away from the target is treated exactly the same as a student who is 20 points away from the target. Similarly, AYP does not recognize that each student has an educational history and performs based on current and past opportunities to learn skills and build knowledge. A test from a student with many opportunities to learn is treated the same way as a test from a student without such advantages. Further, according to AYP, school performance is heavily influenced by the characteristics of the students who enroll in the school rather than how well the school instructs its students. For example, a school that happens to have 20% of its incoming student classified as gifted and talented will have better average performance than a school that has only 5% of its incoming students classified as such. Further, a model classifying schools based on a cut score will not provide a good indicator of school quality (Choi, Goldschmidt, & Yamashiro, 2006, p. 12).

Table 2.1 showcased a comparison in the fundamental differences between status, growth, and value-added models of school accountability.

Table 2.1

Status, Growth, and Value-Added Models

General Considerations by Accountability Model	Status Models	Growth Models	Value-Added Models
Currently approved by ED for NCLB	Yes	No	No
Underlying purpose	Rank/rate schools based on current performance	Rank/rate schools based on performance change	Rank/rate schools based on performance changes different from expected
Major issues for consideration			
Results aligned with AYP	Very likely	Less likely	Less likely
Rating generally understood	Very likely	Likely	Less likely
Inferences same as AYP	Yes	No	No
Requires more than 1 year of data	No	Yes	Yes
Unique student ID required	No	Generally	Generally
Potentially confounds student and school effects	Yes	Less likely	Less likely
Implementation time	Quick	Moderate/varies	High/varies
Implementation process	Simple	Moderate/varies	High/varies
Optimal testing requirements	None	Annual/same content	Annual/same content
Estimate teacher effects	No	Possible	Possible
Possible to measure within school inequities in performance	Limited	Possible	Possible
Costs	Low	Moderate	Potentially high
Simultaneously suitable for program evaluation	Unlikely	Yes	Yes
Measures change for individual students	No	Yes	Yes

Table 2.1 (continued)

General Considerations by Accountability Model	Status Models	Growth Models	Value-Added Models
Absolute	-	Possible	Possible
Relative to Standard	-	Possible	Possible
Requires equal interval scale	Yes	Yes	Yes
Requires vertically equated scale score	No	Varies	Varies
Successful school profile	High average achievement, or exceeds % proficient target	High average achievement growth given average student enrollment	Higher than expected achievement growth given average student enrollment
Intended consequences	Reward high performing school	Rewards growth	Rewards better than expected growth
Unintended consequences	Fosters status quo Ignores within school inequities Rewards schools with "favorable" enrollment Does not reward student achievement growth (school improvement) Reduces incentives for high quality teachers to teach	May ignore high achieving schools May ignore within school inequities Perceived different standards for different sub-groups	May ignore high achieving schools May ignore within school inequities Perceived different standards for different sub-groups

Source: Goldschmidt, P., Roschewski, P., Choi, K., Auty, W., Hebbler, S., Blank, R., & Williams, A. (2005, October). *Policymakers' guide to growth models for school accountability: How do accountability models differ?* Retrieved from http://www.ccsso.org/Documents/2005/Policymakers_Guide_To_Growth_2005.pdf

Throughout the United States, the four most commonly used accountability models are: status, improvement, growth, and value-added models (Goldschmidt et al., 2005). Prior to the enactment of NCLB, some states relied on status-based approaches for the evaluation and measure of state accountability, but after NCLB was fully

implemented, all state-approved accountability systems shifted to status based approaches to determine evaluation and measurement of school performance based on student achievement (Betebenner & Linn, 2009).

Status Model

Goldschmidt et al. (2005) explained that a "status model analyzes school educational achievement compared against an established performance target—usually for one specific school year (p. 3). In other words, a status model focuses on an individual year of assessment results by using them as a predictor of school performance, followed by implementing "decision rules to those results" (Goldschmidt & Choi, 2007, p. 2).

Other researchers explained the model as representative of a snapshot of one moment in time for an assessment performance with proficiency levels compared to a pre-determined target as in AYP goals (Yu, Kennedy, Teddlie, & Crain, 2007; Zvoch & Stevens, 2008).

Further, status models are "often contrasted to growth models. . . . Progress is defined by the percentage of students achieving at the proficient level for that particular year, and the school is evaluated based on whether the student group met or did not meet the goal" (Goldschmidt et al., 2005, p. 3). Essentially, there is one governing question, along with a possible target for schools to meet under the status model, which is, "On average how are student performing this year" (Goldschmidt et al., 2005, p. 3)? Status models, such as the example in Figure 2.4, are focused on how students are performing for a given year with a specific target that must be met by schools.

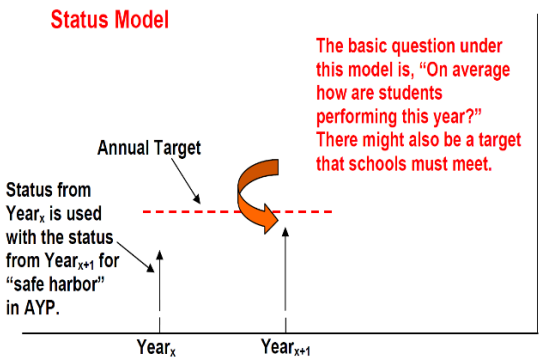


Figure 2.4. Status Model

Source: Goldschmidt, P., Roschewski, P., Choi, K., Auty, W., Hebbler, S., Blank, R., & Williams, A. (2005, October). *Policymakers' guide to growth models for school accountability: How do accountability models differ?* Retrieved from http://www.ccsso.org/Documents/2005/Policymakers_Guide_To_Growth_2005.pdf

Growth Model

Researchers from The Council of Chief State School Officers (CCSS) have acknowledged an increasing interest in the use of growth models in school accountability. Considering that growth models have been used in research and evaluation of program performance for several years, there is a growing trend of interest from policymakers from local, state, and national levels examining the probability for growth models to provide either a replacement or new model that is more helpful to accountability structures in implementation of NCLB (Goldschmidt et al., 2005).

With questions that researchers hold regarding the plausibility of status models, with regards to students' background factors, previous assessment results and the relationship between current student performance in relation to the proficiency targets, the growth model serves as an "alternative to status models" (Goldschmidt & Choi, 2007,

p. 4). Growth models rely on two or more years of performance assessment results as a marker for overall school performance and assigns "decision rules to changes in performance" (Goldschmidt & Choi, 2007, p. 2).

Monitoring AYP in diverse school settings with a larger number of student groups creates more situations of falling short of making AYP progress for demographic categories (Novak & Fuller, 2003). In November of 2005, U.S. Secretary of Education Margaret Spellings announced a Growth Model Pilot program (U.S. Department of Education, 2006) to which states were allowed to submit alternative accountability models to monitor schools (Goldschmidt & Choi, 2007). These growth models offer an alternate way to track student assessment other than NCLB requirements (Goldschmidt & Choi, 2007). The research examined the various types and purposes of accountability models and found that grades and rankings were specific to the measure and kinds of tests used to assess students. For example, in some systems, if a student made academic growth, but the growth was not proportionate to the previous year, it could reflect negatively on the school as not making adequate progress. Other systems focus on how the bulk of students performed in one year compared to the previous year. Interpretation of results depended on the variables of measurement examined.

The researchers involved with the *Guide to Growth Models for School Accountability: How Do Accountability Models Differ?* (Goldschmidt et al., 2005) were in need of additional information about growth models in response to increasing attention and use of growth models for school accountability. Four types of accountability models were created to demonstrate growth, including growth, status, value-added, and improvement. Growth models, as demonstrated in Figure 2.5, focused on how much

students' performance changed on average with an expected growth target to reach. This particular model allows for schools to focus on making progress based on where students begin and where their identified goals are throughout the year. The intended focus is for all students to make some sort of growth. Growth models may benefit schools with student populations with more room for growth and propose a hardship for schools with already higher performing students. Although this study is about the impact of student demographics and school characteristics, growth models may play a part in the outcomes of measuring accountability.

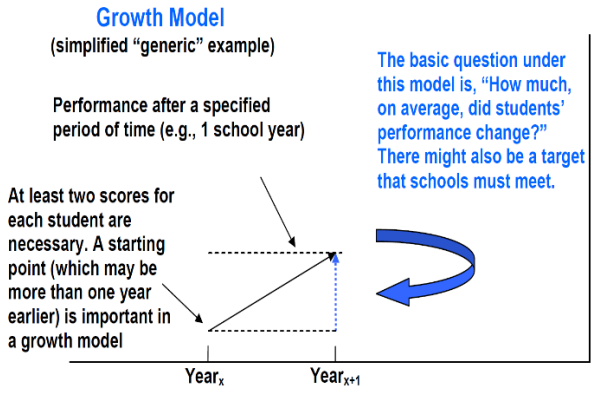


Figure 2.5. Growth Model

Source: Goldschmidt, P., Roschewski, P., Choi, K., Auty, W., Hebbler, S., Blank, R., & Williams, A. (2005, October). *Policymakers' guide to growth models for school accountability: How do accountability models differ?* Retrieved from http://www.ccsso.org/Documents/2005/Policymakers_Guide_To_Growth_2005.pdf

Value-Added Measures

The work compiled by Jaggia and Tuerck (2000) is noteworthy due to their use of statistical models to not only examine various factors affecting student outcomes, but also

to complete a comparison between low-SES schools based on those that exceeded their predictive performance known as *value added* Value-added is a term that indicates methods to measure changes over time in student performance (Ford & Rice, 2015). Value-added systems of accountability that measure student learning instead of student achievement have the probability to better calculate school performance (Goldschmidt & Choi, 2007; Goldschmidt et al., 2005; Lissitz, Doran, Schafer, & Willhoft, 2006; Stevens & Zvoch, 2006). Ready (2012) explained that value-added models created vastly different findings dependent on the model used to measure student progress knowing that achievement and gains were dependent on the lens of the questions being asked and models of assessment used.

> The overarching concern with the phenomena described here is that various model specifications will differentially reward or punish students, teachers, families, and communities. Schools may appear to be making solid academic progress with one approach yet viewed as academically stagnant with an alternative technique. This confusion is unfortunate, as value-added approaches have the potential to highlight the learning that occurs in otherwise low-performing schools, which may provide the positive incentives and recognition needed to attract talented educators to underserved schools and communities. The concern highlighted in this study is that such acknowledgment may be withheld when rightly warranted or rewarded when undeserved. (Ready, 2012, p. 114)

Another problem noted with value-added measures was that underperforming students appeared to make more gains compared to those students who were higher performing (Ready, 2012). In determining accurate accountability for schools, policymakers or

"those responsible for school accountability systems must be aware of how the relationships between initial achievement and subsequent gains influence estimates of school effectiveness and be cognizant of the fact that analytic models that address different questions are likely to produce different results" (Ready, 2012, p. 115).

"Although researchers argue that value-added accountability models allow for identification of achievement and growth, a criticism is the premise that these systems disproportionately penalize socioeconomically disadvantaged schools" (Ready, 2012, p. 115).

Value-added models used for accountability systems have the likelihood to correctly guess what the primary impacts are on student learning, but "student level relationships between initial academic achievement and subsequent academic growth may reduce the validity of such approaches" (Ready, 2012, p. 98). Value-added models observed whether students changed in performance and met their growth targets along with how much the student met, missed, or exceeded growth expectations. For example, Figure 2.6 shows the premise behind value-added systems.

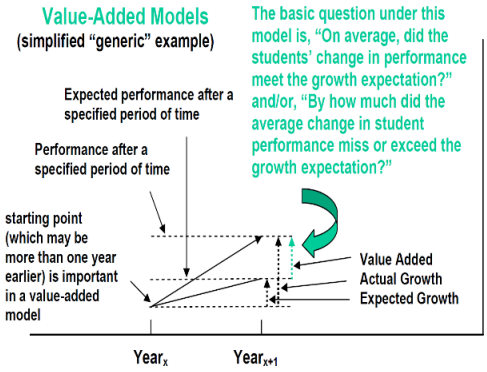


Figure 2.6. Value-Added Model

Source: Goldschmidt, P., Roschewski, P., Choi, K., Auty, W., Hebbler, S., Blank, R., & Williams, A. (2005, October). *Policymakers' guide to growth models for school*

accountability: How do accountability models differ? Retrieved from http://www.ccsso.org/Documents/2005/Policymakers_Guide_To_Growth_2005.pdf

Improvement Model

Improvement models examine a student's progress in relation to where a different group of students were in performance the prior year at the same time. One item of concern with this model is that the student group is measured compared to an entirely different group of students with possibly different teachers from the year before. Improvement of students during one year of instruction cannot predict or inform how students will do another year. For example, Figure 2.7 shows the premise behind improvement model systems.

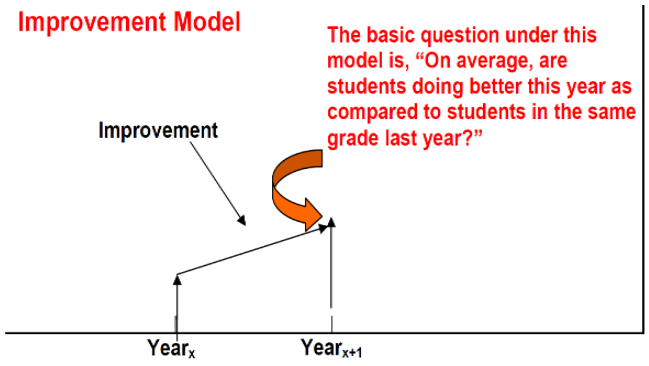


Figure 2.7. Improvement Model

Source: Goldschmidt, P., Roschewski, P., Choi, K., Auty, W., Hebbler, S., Blank, R., & Williams, A. (2005, October). *Policymakers' guide to growth models for school accountability: How do accountability models differ?* Retrieved from http://www.ccsso.org/Documents/2005/Policymakers_Guide_To_Growth_2005.pdf

A problematic issue with school accountability as a fair measure to report school progress are models that monitor a single year performance, without taking into consideration past years (Goldschmidt & Choi, 2007). Novak and Fuller (2003) conveyed that accountability formulas and growth models that were used to assess students over time led to concerns about fairness in comparing groups of students to the performance of students who previously tested the year before in the same grade level, considering the transiency of students and staff teaching students.

Analyses for growth models provide additional data for state education agencies already covered up in data that tell numerous stories, unless researchers provide a more systemic approach to using growth model data to report accountability in a relevant way (Betebenner & Linn, 2009). Braun (2008) advocated for a detailed plan of action, "specifying what data is necessary and how that data will be used that in turn leads to actions with desired outcomes" (p. 20). Three reporting features of accountability influence growth with the greatest impact, including "measurement, longitudinal, data analysis, and accountability." Accountability "carries the greatest weight" (Betebenner & Linn, 2009, p. 20).

A hindrance to the use of performance standards to measure growth lies in the reporting with only three to four levels, which "contain a wide range of achievement and hence have the potential to mask substantial student growth" (Betebenner & Linn, 2009, p. 5). A second limitation is that a student could score proficiently one year and then achieve the same score again the next year, which would not report the growth of the student over time. Reporting a category of proficiency each year illustrates a student who maintains the proficient score with new content, but it does not indicate any improvement

scores. Another problem with school grades based on student performance is the differences among institutions from state to state, making the case for a common discussion of standards and measurements from state to state for more accuracy and validity (Betebenner & Linn, 2009).

Quantifying growth vertically from grade level to the next is a common approach used for scale scores, but it is not easily understood in layman's terms (Ballou, 2008; Briggs & Betebenner, 2009). Betebenner and Linn (2009) asserted that student growth data should focus on the quality of data by placing a greater emphasis on the development of models and practices that are useful and practical, while questioning who the data is for and how will it be used (Linn, 2000). These negative concerns are not new and are termed as *Campbell's Law* (Campbell, 2010). The drive toward increasing data availability and quality are positive steps, but subject to the many well-understood and often-recited lessons about the use of large-scale assessment data for high-stakes purposes (Linn, 2000; Mintrop & Sunderman, 2009).

Value-added models are a means for which student performance can be measured over time (Ford & Rice, 2015) and they are encouraged nationally and on the federal level. Some of the concerns with a value-added model are that "these methods are very complex and highly technical . . . [and] they may be used inappropriately" (Condie, Lefgren, & Sims, 2014) or leave out factors that affect the performance of students (Ford & Rice, 2015). Ford and Rice (2015) further explained their research:

> A very basic value-added model was described and applied to the CST data. In this model no student characteristics were controlled. We avoided any distortion to the value-added calculations and did not account for variables such as low SES

densities, or other socioeconomic or racial or any other factors. The underlying assumption in the selection of such a simple model is the idea that all students have similar capacities to learn. Despite our choice, we do not rule out that the control of some variables might need to occur to properly understand the relative quality of program outcomes. The selection of which variables to control requires significant discussion, well beyond the scope of this discussion. With the use of this model, the subsequent value-added results indicated the existence of distinguished online schools that perform above average with 95% statistical confidence in seven of the eight course categories analyzed. (p. 418)

For example, Figure 2.8 shows a grid of four groups, separated into quadrants, comparing degrees of growth and status.

High Growth ***Group III*** *Low Status*	*High Growth* ***Group IV*** *High Status*
Low Growth ***Group I*** *Low Status*	*Low Growth* ***Group II*** *High Status*

Figure 2.8. Growth vs. Status Achievement Grid

Source: Goldschmidt, P., Roschewski, P., Choi, K., Auty, W., Hebbler, S., Blank, R., & Williams, A. (2005, October). *Policymakers' guide to growth models for school accountability: How do accountability models differ?* Retrieved from http://www.ccsso.org/Documents/2005/Policymakers_Guide_To_Growth_2005.pdf

Accountability models are not a new phenomenon and questions remain as to who or what is behind the changes in accountability systems, as well as observations as to who benefits as a result of changes to assessments measuring school performance (Dorn & Ydesen, 2014). Dorn and Ydesen (2014) noted that a different "path is intimately connected with the issue of accountability practices being connected with questions of power, education access, education management, and social selection" (p. 6).

Despite grades for school performance, student groups with language barriers find ways to perform well, showcasing growth in programs through unconventional programs aimed at helping both the students and families of students (Colvin & Helfand, 2000). Another spectacle is the competitive nature of families basing enrollment decisions about information generated by school accountability formulas that are neither neutral nor accurate based on school rankings, ratings, and grades (Glynn & Waldeck, 2013).

Stakeholders from schools in England have more readily accepted multiple measures of school performance through value-added formulas, while the United States is focused solely on test scores for grading schools (Goldschmidt & Choi, 2007). This narrow focus has unintended consequences (Stecher, 2006) with schools focused on limited curriculum and test-prep. Researchers additionally found that "report cards do not have a significant influence on performance. The point estimates, although positive, are not significantly different from zero" (Hanushek & Raymond, 2005, p. 310).

Assignment of Grades and Ranking in the United States

In 2015, the Texas Senate passed legislation as part of Senate Bill 6 (LegiScan, n.d.) sponsored by the Senate Education Committee Chair enacting a school performance rating of A-F, allowing parents and members of communities access to information about

schools (Taylor, 2015). According to Executive Director, Patricia Levesque from Excellence in Education National, "The purpose of a school accountability system is to increase student performance by providing clear and consumable information to parents, educators and the public about school effectiveness" (Taylor, 2015). Communicating a letter grade of A, B, C, D, and F as performance indicators, ensures state constituents have an understanding of whether or not the school is performing well, needs improvement, or is failing in overall academic performance (Taylor, 2015). An A-F rating logically signifies how schools are doing with respect to student learning and quickly identifies whether or not schools are failing or excelling in student achievement (Taylor, 2015).

Pertaining to measures of student performance for state assessments, Toutkoushian and Curtis (2005) stated, "policymakers should be concerned that comparisons of schools that are based on outcome measures do not consider the existence of factors that are correlated highly with these outcomes, many of which are beyond the control of the school" (p. 260).

Assigning Grades in Utah's Schools

Since the inception of grades assigned to schools throughout Utah, beginning in 2011, each year has yielded changes from the legislative sessions (Jacobsen, 2014).

> With 2014 being a baseline year for SAGE results, next year is expected to show improved scores as teachers and students become familiar with the test and the new standard. The Legislature approved a one-time adjustment to the school grading system in light of the transition to SAGE from the CRT, the state's former year-end assessment. (Jacobsen, 2014)

Individual schools in Utah receive a grade determined by a point system that measures overall academic growth, proficiency in math, English and science. High school grades also include graduation rates and ACT scores. All grades are based on Student Assessment of Growth and Excellence (SAGE), which is an assessment system designed to raise the rigor for performance in math, English and science (Jacobsen, 2014).

There are a few issues that negatively affect the grades of schools or might interfere with the study: a) alternative schools are exempt from school grading; b) new schools are allowed to apply for a temporary exemption; c) students who take an alternative assessment are not counted in SAGE, so they potentially may not be counted in the 95% required of the participation rate; and d) if there's a mandatory ranking applied to all schools across Utah and any grade is off, that could have a negative effect on other schools' grades (Jacobsen, 2014).

How School Grades Are Calculated in Utah

Utah's school grading accountability system, known as Grading Utah Schools (GUS), was created to build a transparent and easy way to navigate the evaluation of Utah schools by assigning each school a grade of A, B, C, D or F (Utah State Office of Ed, 2015). One of the board-adopted policies requires "effective assessment to inform high quality instruction and accountability" as part of the state's transition to Utah's college and career-ready standards (Utah State Office of Ed, 2015, p. 11).

School grades are determined by how many points a school obtains from indicators on countable test participants who took viable tests. There is a total of 900 points for high schools (2015, p. 11).

Table 2.2 indicates a total of 900 points for high schools.

Table 2.2

High Schools Grading Scale

Percent of Points	Points	Grade
64% - 100%	572 - 900	A
51% - 63%	455 - 571	B
43% - 50%	383 - 454	C
40 - 42%	356 -382	D
< 40%	≤ 355	F

Source: Utah State Office of Education. (2015, October 1). *SAGE testing and data reporting*. Retrieved from http://www.schools.utah.gov/assessment/SAGE/SAGETestingDataReporting.aspx

Figure 2.9 provides the metrics for assessed categories along with the numeric point values on the SAGE assessment.

- Academic Proficiency (300 points)
- Academic Growth (300 points)
  - Growth of All Students (AS; 150 points)
  - Growth of Below Proficient Students (BPS; 150 points)
- College and Career Readiness (CCR; 300points)
  - Graduation Rate (150 points)
  - ACT Achievement (150 points)
- Participation Rate (Required 95%)

Figure 2.9 Utah's School Grading Metrics

Source: Utah State Office of Education. (2015, October 1). *SAGE testing and data reporting*. Retrieved from http://www.schools.utah.gov/assessment/SAGE/SAGETestingDataReporting.aspx

In March of 2011, state lawmakers, educational leaders, and stakeholders across Utah created Student Federal Accountability Reporting (SFAR), which integrated seven

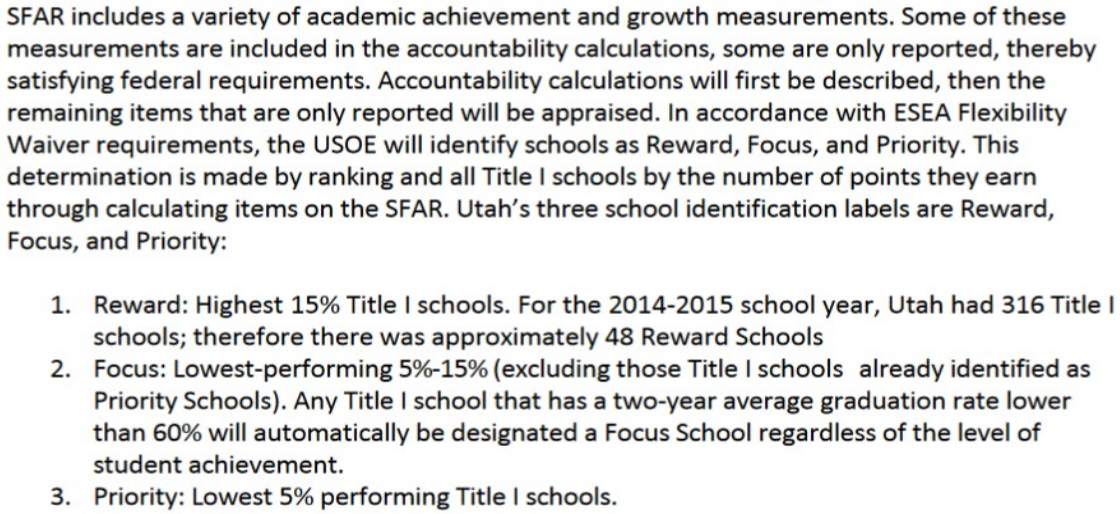
principles into their design, as indicated by Figure 2.10. Built into the framework were multiple achievement and growth measurements.

1. Promote progress toward and achievement of college and career readiness
2. Value both meeting standards (proficiency) and improving academic achievement (growth)
3. All schools, including those that serve traditionally low performing students, should have an opportunity to demonstrate success
4. Strong incentives for schools to improve achievement for the lowest performing students
5. Growth expectations for non-proficient students should be linked to attaining proficiency
6. Growth expectations for all students, including students above proficiency, should be appropriately challenging and meaningful
7. Clear and understandable to stakeholders

Figure 2.10. Student Federal Accountability Reporting Design Principles

Source: Utah State Office of Education. (2015, October 1). *SAGE testing and data reporting*. Retrieved from http://www.schools.utah.gov/assessment/SAGE/SAGETestingDataReporting.aspx

Figure 2.11 provides an overview of the Student Federal Accountability Reporting (SFAR) with a breakdown of the reward, focus and priority categories.



SFAR includes a variety of academic achievement and growth measurements. Some of these measurements are included in the accountability calculations, some are only reported, thereby satisfying federal requirements. Accountability calculations will first be described, then the remaining items that are only reported will be appraised. In accordance with ESEA Flexibility Waiver requirements, the USOE will identify schools as Reward, Focus, and Priority. This determination is made by ranking and all Title I schools by the number of points they earn through calculating items on the SFAR. Utah's three school identification labels are Reward, Focus, and Priority:

1. Reward: Highest 15% Title I schools. For the 2014-2015 school year, Utah had 316 Title I schools; therefore there was approximately 48 Reward Schools
2. Focus: Lowest-performing 5%-15% (excluding those Title I schools already identified as Priority Schools). Any Title I school that has a two-year average graduation rate lower than 60% will automatically be designated a Focus School regardless of the level of student achievement.
3. Priority: Lowest 5% performing Title I schools.

Figure 2.11. Student Federal Accountability Reporting Academic Achievement and Growth Measurements

Source: Utah State Office of Education. (2015, October 1). *SAGE testing and data reporting*. Retrieved from http://www.schools.utah.gov/assessment/SAGE/SAGETestingDataReporting.aspx

Figure 2.12 provides a visual for how both proficiency and growth are calculated from the number of percent proficient for *all students* plus *graduation rate* and then tabulates growth based on the number of students who made growth out of the *all students* category and *below proficient students* category. For high schools, 300 points are divided in half with 150 from the percent reaching proficiency and half based on graduation rate. (Utah State Office of Ed, 2015).

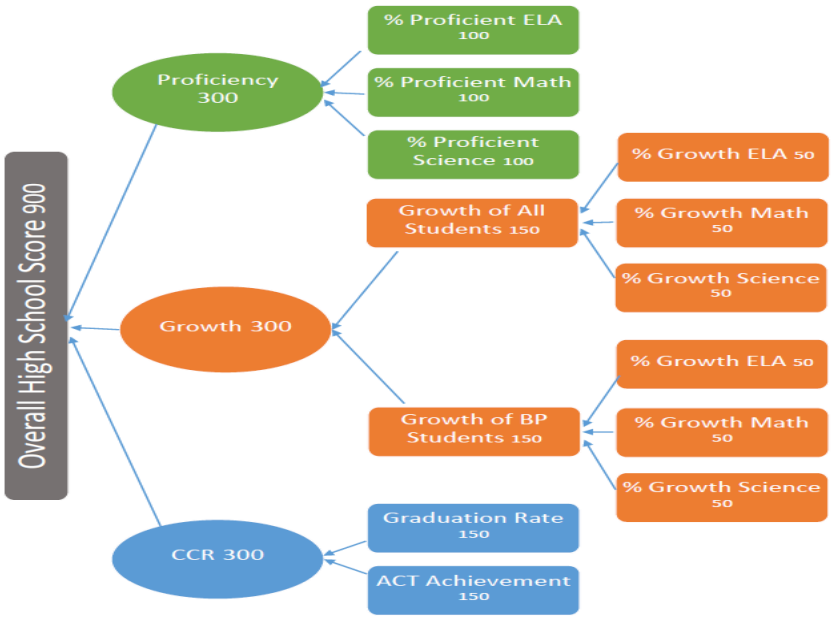


Figure 2.12. School Grading Points for High Schools

Source: Utah State Office of Education. (2015, October 1). *SAGE testing and data reporting*. Retrieved from http://www.schools.utah.gov/assessment/SAGE/SAGETestingDataReporting.aspx

Table 2.3 provides one example of a hypothetical high school calculation for English Language Arts, mathematics, and science tests from the Student Federal Accountability Reporting (SFAR). These tests are used to determine percentages of students who reached proficient out of the total number of students tested along with the impact of that number on overall points awarded.

Table 2.3

Example of SFAR/PACE Proficiency Calculation for High Schools

SAGE Test	Number of Proficient Scores from Countable Participants	Total Number of Scores from Countable Participants	Percent Proficient	Points Possible	Test Earned Points
ELA	25	100	25.00%	50	13
Math	63	77	81.82%	50	41
Science	20	32	62.50%	50	31
Total Proficiency Points					85

Source: Utah State Office of Education. (2015, October 1). *SAGE testing and data reporting*. Retrieved from http://www.schools.utah.gov/assessment/SAGE/SAGETestingDataReporting.aspx

Graduation Rate

The graduation rate for SFAR/PACE accounts for 150 of the 300 points towards achievement. The current graduation rate rules come from No Child Left Behind (NCLB) guidance. The following formula provides an example of the four-year graduation rate for the cohort entering ninth grade for the first time in the fall for the 2011-2012 school year (Utah State Office of Ed, 2015).

Figure 2.13 provides an overview of graduation rate criteria, as well as a breakdown for how points are used to determine growth for *all students* and *below proficient students*.

Number of cohort members who earned a regular high school diploma by the end of summer 2015

Number of first-time 9th graders in fall 2011 (starting cohort) plus students who transfer in, minus students who transfer out, emigrate, or die during school years 2011-2012, 2012-2013, 2013-2014, 2014-2015

Other guidelines for graduation rate are:

1. If a student graduates prior to or when their cohort graduates they are considered a graduate.
2. If a student graduates after the summer of their 12th grade year, or never graduates, they are counted as a non-graduate.
3. The last school that a student enrolls in is accountable for the student's graduation status. However, if a student attends two schools in their final year, and one school graduates the student while the other does not, then the school graduating the student is held accountable. If neither school graduates the student, then the school with the latest exit date is accountable.
4. The graduation rate portion of the Achievement calculation is removed for high schools with 10 or fewer graduates. This means that a high school with fewer than 10 graduates will only have 150 points possible Achievement.
5. In the rare case where a high school has viable SAGE scores but no graduation rate, then the graduation rate is removed from the calculation and the 300 total possible points for Academic Achievement is only based solely on academic proficiency based on the SAGE proficiency percentages.
6. In the other rare case where a high school has a graduation rate, but no viable SAGE scores Academic Proficiency is solely based on the graduation rate.

Growth

A total of 300 points is possible for Growth (200 points for AS; 100 points for BPS). The Student Growth Percentile (SGP) method provides the basis for measuring academic growth in the model. Student growth is determined by comparing each student's progress with that of other students in the state with the same prior achievement pattern. SGPs provide a familiar basis to interpret performance, the percentile, which indicates the probability of an outcome given the student's starting point and can be used to gauge whether the student's growth was atypically high or low.

Figure 2.13. Other Guidelines for Graduation Rate/Growth Formula

Source: Utah State Office of Education. (2015, October 1). *SAGE testing and data reporting*. Retrieved from http://www.schools.utah.gov/assessment/SAGE/SAGETestingDataReporting.aspx

Like Utah's school grades, student growth percentiles (SGPs) are calculated for all countable students with a minimum of two viable SAGE scores in a given content area. Growth is evaluated in the same way for all schools (elementary, middle, and high

schools). For SFAR/PACE, there are three levels of growth based on median growth percentile (Utah State Office of Ed, 2015).

Table 2.4 provides a rubric for evaluating median growth percentiles by *all students* and *below proficient students*.

Table 2.4

Rubric for Evaluating Median Growth Percentiles by Group

MGP Ranges	All Students	Below Proficient Students
≥ 70	200	100
30-69	35	50
< 30	(MGP x 3.75) – 62.50	(MGP x 1.875) – 31.25

Source: Utah State Office of Education. (2015, October 1). *SAGE testing and data reporting*. Retrieved from http://www.schools.utah.gov/assessment/SAGE/SAGETestingDataReporting.aspx

The rubric is used for each of the three SAGE content areas (ELA, math, and science) evaluated. The average of all of the test MGPs are used to calculate a composite growth measure for all students (AS) and below proficiency students (BPS). Each test is weighted equally. Other guidelines for SFAR growth calculations include the following:

1. If there are fewer than 10 students in a subject area, do not calculate a score for that subject area.
2. If there are fewer than 10 students in each of the three subject areas, do not calculate for that group.
3. If there are fewer than 10 students in each of the three subject areas of the BPS, multiply each of the AS subject's points by 1.5.

4. If there are fewer than 10 students in each of the three subject areas, do not calculate BPS Growth. Possible points will transfer to AS Growth (Utah State Office of Ed, 2015).

Research Question

What extent do student demographics and school characteristics predict school accountability grades? The research question examines whether student demographics and school characteristics impact grades. Student characteristics include low SES students, racial minority, English Language Learners, and disability. School characteristics include teachers with a graduate degree and rural location.

CHAPTER 3: METHODS

Introduction

This study addressed gaps in the literature to identify predictors of variance within student and school demographics in order to determine the implications of these predictors on state accountability systems and the grades they assign to schools. There is a particular need to determine if student demographics, combined with school characteristics, impact state ranking and grading systems with regards to equity. Across the United States, schools are ranked according to student and school performance, ranging from high performing to low performing with "rewards, recognition, and consequences" connected to each of these categories (Traylor, 2013, p. 42). Identifying trends between student demographics and/or school characteristics within school accountability rankings is necessary in order to determine whether or not accountability systems of integrity are responsible for building instructional capacity or conversely, are deepening the divide between the *haves* and *have nots*. If student and school variables can predict state accountability results, grades and rankings could be contributing to a social justice problem ultimately negatively impacting student achievement.

Research Question and Hypothesis

To what extent do student demographics and school characteristics predict school accountability grades or rankings? The research question examines whether student demographics and school characteristics impact school accountability grades and rankings. This study utilized student demographic categories for predictors such as low SES, English Language Learners, disability, and racial minorities on school

accountability grades, along with school characteristics, including teachers with a graduate degree and rural locations.

Description of Research Design

This study will examine the predictive power of student demographics and school characteristics on GUS grade, GUS overall points earned, graduation rate, ACT score, and achievement sub-scores. Analyses from the Student Assessment of Growth and Excellence (SAGE) were conducted using statewide testing data from the Utah State Office of Education. Six dependent variables served as indicators for student demographics and school characteristics collected from the Utah State Office of Education database. The independent variables for student demographics are low SES, English Language Learners, disability, and racial minority. Additional independent variables representing school characteristics are the number of teachers with a graduate degree and significance of performance within rural locations.

The dependent variables in this study are GUS grade, GUS overall points earned, graduation rate, ACT score, and achievement sub-scores.

Comparative and multiple regression test analyses (Field, 2009; Stevens, 1999) are proposed along with calculating bivariate correlations (Field, 2009; Stevens, 1999) in order to evaluate all relationships between overall school rankings with each independent variable (George & Mallery, 2011). By conducting both descriptive statistics (Stevens, 1999) and correlational analyses (Field, 2009), the research allowed for additional insights on the relationships between diverse students and school populations and school accountability grades. These findings indicated the extent to which those variables collectively explain the variance (Field, 2009; Stevens, 1999) in school accountability

grades. The implications, impact, and influences of student demographics and student characteristics on school accountability will be reported. Through running multiple regression analyses (Field, 2009; Stevens, 1999), the researcher determined significant predictors of the dependent variables.

A multiple regression analysis (Field, 2009; Stevens, 1999) allows more than one independent variable to have an influence on the dependent variable (George & Mallery, 2011, p. 92). The researcher attempted to predict the grading based on student demographics and school characteristics.

Population Data

The data used for analyses included high schools (grades 9-11) tested in reading and math using the Student Assessment of Growth and Excellence (SAGE) during the 2014-2015 assessment window. The Utah schools data reflect approximately 35,000 students in 361 high schools, which were used for comparative and multiple regression analyses (Field, 2009; Stevens, 1999).

Sample

A sample ($n = 361$) of high schools across the state of Utah were used to investigate the research question. The comparison and multiple regression analysis (Field, 2009; Stevens, 1999) test data revealed statistically significant ($p < 0.01$) relationship between the student demographic categories and school characteristics and to the school accountability rankings in Utah. This analysis demonstrated the extent to which school demographics or school characteristics negatively or positively affect accountability rankings.

Instrumentation

Utah uses a combined reading and math test for state assessment and accountability known as Student Assessment of Growth and Excellence (SAGE). A committee comprised of Utah public school teachers developed the computer-adaptive assessments in which tests are designed to allow students to move to more advanced or less difficult questions based on the amount of correct or incorrect answers. All test items were reviewed by a 15-member parent committee with at least two parents assigned to review each test question (Utah State Office of Education Assessment and Accountability, n.d.). The SAGE assessments are given annually and reports provide both the scaled score and proficiency level for each assessment taken by students. The scaled score indicates the student's performance, which is converted to a common scale (Utah State Office of Education Assessment and Accountability, n.d.). Figure 3.1 provides an explanation as to how SAGE is reported with regard to scaled score and proficiency level.

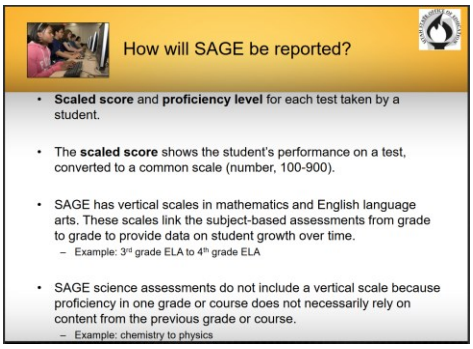


Figure 3.1. SAGE Reporting I

Source: Utah State Office of Education Assessment and Accountability. (n.d.). *Preparing Utah's students for college and career: New standards, new tests, new scores*.

Retrieved from

http://schools.utah.gov/assessment/SAGE/StandardsTestScores.aspx

Utah's State Office of Education has taken steps to institute measures to interpret scale scores on a scale from 100-900 to determine proficiency by separating scores into four categories: *highly proficient*, *proficient*, *approaching proficient*, and *below proficient*. Students who score *highly proficient* or *proficient* are on track for college and career readiness, while those scoring *below proficient* and *proficient* are not. (Utah State Office of Education Assessment and Accountability, n.d.). Figure 3.2 illustrates additional information on how SAGE is reported with scaled scores and proficiency levels.

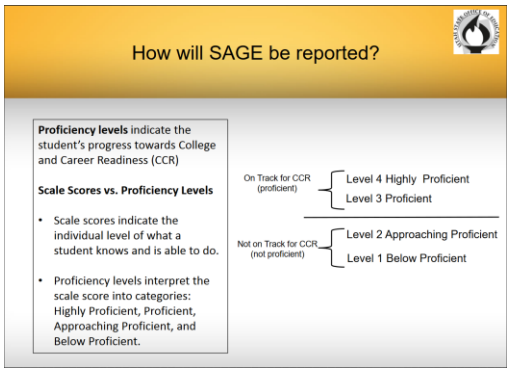


Figure 3.2. SAGE Reporting II

Source: Utah State Office of Education Assessment and Accountability. (n.d.). *Preparing Utah's students for college and career: New standards, new tests, new scores*.

Retrieved from

http://schools.utah.gov/assessment/SAGE/StandardsTestScores.aspx

Data Collection and Analysis

Descriptive and correlational statistics (Field, 2009) were utilized from multiple databases made available to the researcher by the Utah State Office of Education.

Microsoft Excel spreadsheets containing overall SAGE scores for reading and math, along with school characteristics, were downloaded from the Utah State Office of Education's database. All files were identified by an individual school identification number, allowing all files to be combined into one SPSS file in order to determine positive or negative effects on variables used to determine school accountability grades. Approval for research was obtained by Eastern Kentucky University Institutional Review Board.

CHAPTER 4: RESULTS

Descriptive Statistics

Descriptive statistics (Stevens, 1999) were analyzed to determine similarities between participants. Table 4.1 illustrates the means and standard deviations for six predictors of variance. In the 361 high schools, the mean percentage of low SES students was 39.9. This is the highest percentage of any student characteristic sub-group. The lowest mean percentage of any student characteristic sub-group was English Language Learners at 3.71. The mean percentage of racial minority students was 15.31. Considering the low percentage of minorities in the state out of the whole, this number is not surprising. The percentage of teachers with a graduate degree in participating schools was 44.09, almost half of all teachers.

Table 4.1

Mean Percent for High Schools

	N	Minimum	Maximum	Mean	Std. Deviation
% Low Income	361	.00	100.0	39.85	21.38
% English Language Learners	361	.00	73.00	3.71	7.17
% Students with Disabilities	361	.00	100.00	13.58	13.36
% Racial Minority	361	.00	88.00	15.31	12.87
% Teachers with a Graduate Degree	361	.00	100.00	44.09	19.63

Table 4.2 provides an overview of the number of rural schools observed in Utah.

Table 4.2

Percentage of Rural Schools

		Frequency	Valid Percent	Cumulative Percent
Valid	No	275	76.2	76.2
	Rural	86	23.8	100.0

Correlational Analysis

Bivariate correlations (Field, 2009) of the independent variables showed relationships that were expected. For example, there was a moderate positive correlation (Field, 2009) between percent low SES and percent ELLs (r = .54) as well as a moderate positive correlation (Field, 2009) between percent low SES and percent racial minority (r = .47). All other independent variables showed positive bivariate correlations with all independent variables, but none as strong as those between low SES, ELL and race.

Table 4.3

Intercorrelation Matrix

		% Low Income	% ELLs	% Students with Disabilities	% Racial Minority	% Teacher with Graduate degree	Rural School
% Low Income	Pearson Correlation	1	.54	.29	.47**	.05	.11*
	Sig. (2-tailed)		.00	.00	.00	.33	.03
	N	361	361	361	361	361	361

Table 4.3 (continued)

		% Low Income	% ELLs	% Students with Disabilities	% Racial Minority	% Teacher with Graduate degree	Rural School
% Students with Disabilities	Pearson Correlation	.29**	.04	1	.05	-.02	.02
	Sig. (2-tailed)	.00	.41		.31	.65	.76
	N	361	361	361	361	361	361
% Racial Minority	Pearson Correlation	.47**	.38**	.05	1	.04	-.31
	Sig. (2-tailed)	.00	.00	.31		.41	.00
	N	361	361	361	361	361	361
% Teachers with Graduate Degree	Pearson Correlation	.05	-.02	-.02	.04	1	-.14**
	Sig. (2-tailed)	.33	.66	.65	.41		.01
	N	361	361	361	361	361	361
Rural School	Pearson Correlation	.11*	-.02	.02	-.31**	-.14**	1
	Sig. (2-tailed)	.03	.75	.76	.00	.01	
	N	361	361	361	361	361	361

*Correlation is significant at the 0.05 level (2-tailed)

**Correlation is significant at the 0.01 level (2-tailed)

Multiple Regression Analysis

A standard multiple regression analysis (Field, 2009) was performed between student and school characteristics (low SES, ELL, students with disabilities, racial minority, and teachers with a graduate degree) and GUS school accountability grade. Regression analysis revealed that the model significantly predicted school accountability grade, $F(6, 276) = 23$, $p < .01$. Adjusted R^2 for the model was .32. This indicates that 32% of the variance in school accountability grade was accounted for by the variables. Although this is a moderate effect size (Field, 2009), it is still significant. Low income ($\beta = -.30$, $t = -3.95$, $p < .01$), ELL ($\beta = -.27$, $t = -3.92$, $p < .01$) and percent of teachers with a graduate degree ($\beta = .14$, $t = 2.8$, $p < .01$) each significantly predicted school accountability grade. All of the independent variables used in this research are shown in Table 4.4. The significance determined the impact on the grade earned.

Table 4.4

School Accountability Grade and Student and School Characteristics

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	3.87	.18		21.36	.00
	% Low Income	-.02	.00	-.30	-3.97	.00
	% ELLs	-.04	.01	-.27	-3.92	.00
	% Students with Disabilities	.00	.01	.04	.75	.46
	% Racial Minority	-.00	.01	-.06	-.85	.40
	% Teachers with Graduate Degree	.01	.00	.14	2.80	.01

Table 4.4 (continued)

Model	Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
Rural School	-.08	.13	-.03	-.58	.56
a. Dependent Variable: Grade Earned					

A standard multiple regression analysis (Field, 2009) was also performed between student and school characteristics (low SES, ELL, students with disabilities, racial minority and teachers with graduate degree) and GUS overall points earned. Regression analysis revealed that the model significantly predicted overall points earned, $F(6, 230) = 22.4$, $p < .01$. Adjusted R^2 for the model was .36. This indicates that 36% of the variance in overall points earned was accounted for by the variables. Although this is a moderate effect size (Field, 2009), it is still significant. ELLs ($\beta = -.28$, $t = -3.70$, $p < .01$), students with disabilities ($\beta = -.27$, $t = -4.42$, $p < .01$) and percent of teachers with a graduate degree ($\beta = .18$, $t = 3.28$, $p < .01$) each significantly predicted overall points earned. Table 4.5 provides an overview of all of the independent variables used within this research to determine overall points earned.

Table 4.5

Grading Utah Schools Overall Points Earned

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	345.82	14.69		23.54	.05
	% Low Income	-.51	.26	-.17	-2.00	.00
	% ELLs	-2.39	.65	-.28	-3.68	..00

Table 4.5 (continued)

Model	Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
% Students with Disabilities	-4.64	1.05	-.27	-4.43	.00
% Racial Minority	-.12	.30	-.03	-.38	.70
% Teachers with Graduate Degree	.61	.19	.18	3.30	.00
Rural School	3.77	8.64	.03	.44	.66

a. Dependent Variable: GUS Overall Points

Additionally, a standard multiple regression analysis (Field, 2009) was performed between student and school characteristics and GUS sub-scores. The sub-scores analyzed were: a) Growth in English Language Arts: All students; b) Growth in English Language Arts: Below proficient students; c) Growth in Math: All students; d) Growth in Math: Below proficient students; e) Growth in Science: All students; f) Growth in Science: Below proficient students; g) English Language Arts proficient; h) Math proficient; i) Science proficient; j) College & career readiness graduation rate; and k) College & career readiness ACT.

Regression analysis revealed that all sub-scores significantly predicted overall points earned, $F(6, 210) = 87$, $p < .01$. Adjusted R^2 for the model was .82. This is a strong effect size (Field, 2009) when combining all student and school characteristics to predict variance, considering only two of the sub-scores, English Language Arts Proficient ($\beta = .28$, $t = 3.01$, $p < .01$) and College and Career Readiness Graduation Rate ($\beta = .20$, $t = 5.54$, $p < .01$) significantly predicted points earned for Grading Utah Schools.

In Table 4.6, all of the independent variables are listed for the Grading Utah's School sub-scores to look for predictors of variance.

Table 4.6

Grading Utah Schools Sub-Scores

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-2.40	.34		-7.11	.00
	Growth in English Language Arts: All students	.02	.01	.12	1.84	.07
	Growth in English Language Arts: Below proficient students	-.00	.01	-.01	-.22	.83
	Growth in Math: All students	.022	.01	.13	2.12	.04
	Growth in Math: Below proficient students	-.00	.01	-.01	-.16	.88
	Growth in Science: All students	.01	.01	.04	.47	.64
	Growth in Science: Below proficient students	.02	.01	.15	2.21	.03
	English Language Arts proficient	.02	.01	.28	3.01	.00
	Math proficient		.00	.13	1.95	.05
	Science proficient		.01	.15	2.02	.05
	College & career readiness graduation rate		.00	.20	5.54	.00
	College & career readiness ACT		.00	.12	2.29	.02

a. Dependent Variable: Grade Earned

For the English Language Arts Proficiency sub-score, the student characteristics percent low SES (t = -6.45, p < .01), percent English Language Learners (t = -3.40, p < .01) and percent students with disabilities (t =3, p < .01) significantly predicted (Field, 2009) points earned for Grading Utah Schools

In Table 4.7, all of the independent variables were compared to the Grading Utah Schools' English Language Arts Proficiency Sub-Scores.

Table 4.7

Grading Utah Schools English Language Arts Proficiency Sub-Scores

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	49.87	2.41		20.72	.00
	% Low Income	-.35	.05	-.47	-6.45	.00
	% ELLs	-.49	.15	-.23	-3.40	.00
	% Students with Disabilities	.24	.08	.15	2.98	.00
	% Racial Minority	.06	.07	.06	.94	.35
	% Teachers with Graduate Degree	.08	.04	.11	2.24	.03
	Rural School	-.24	1.76	-.01	-.14	.90

a. Dependent Variable: English Language Arts Proficient

A standard multiple regression analysis was then performed between student and school characteristics (low SES, ELL, students with disabilities, racial minority, teachers with graduate degree, rural school location) and graduation rate. Regression analysis revealed that the model significantly predicted overall points earned, $F(6, 253) = 7.14$, $p <. 01$. Adjusted R^2 for the model was .13. This indicates that 13% of the variance in graduation rate was accounted for by the variables. Although this is a small effect size (Field, 2009), it is still significant. ELLs ($\beta = -3$, $t = -3.26$, $p < .01$) and percent of teachers with a graduate degree ($\beta = .21$, $t = 3.43$, $p < .01$) each significantly predicted graduation rates.

In Table 4.8, all of the independent variables were compared to the Grading Utah Schools' graduation rate.

Table 4.8

Graduation Rate

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	128.20	3.80		33.77	.00
	% Low Income	.05	.08	.06	.60	.55
	% ELLs	-.77	.24	-.30	-3.26	.00
	% Students with Disabilities	-.20	.27	-.05	-.74	.46
	% Racial Minority	-.10	.09	-.09	-1.17	.24
	% Teachers with Graduate Degree	.17	.05	.21	3.43	.00
	Rural School	4.74	2.43	.14	1.95	.05

a. Dependent Variable: College and Career Readiness Graduation Rate

A standard multiple regression analysis was finally performed between student and school characteristics (low SES, ELL, students with disabilities, racial minority, teachers with graduate degree, rural school location) and College and Career Readiness ACT scores. Regression analysis revealed that the model significantly predicted overall points earned, $F(6, 253) = 40.72$, $p < .01$. Adjusted R^2 for the model was .49. This indicates that 49% of the variance in school accountability grade was accounted for by the variables. This is a large effect size (Field, 2009), which is significant. Low income ($\beta = -.27$, $t = -3.55$, $p < .01$) and students with disabilities ($\beta = -.48$, $t = -9.1$, $p <. 01$) each significantly predicted ACT scores.

In Table 4.9, all of the independent variables were compared to the Grading Utah Schools' college and career readiness ACT scores.

Table 4.9

College and Career Readiness ACT Scores

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	78.05	3.86		20.23	.00
	% Low Income	-.27	.08	-.27	-3.56	.00
	% ELLs	-.39	.24	-.11	-1.62	.11
	% Students with Disabilities	-2.48	.27	-.48	-9.10	.00
	% Racial Minority	.05	.09	.04	.59	.56
	% Teachers with Graduate Degree	-.04	.05	-.04	-.83	.41
	Rural School	-2.39	2.47	-.05	-.97	.34

a. Dependent Variable: College and Career Readiness ACT

Running descriptive statistics (Stevens, 1999), correlational analyses (Field 2009), and multiple regression analyses (Field, 2009) provided insight into which predictors of variance had the strongest correlation to predicting school grades. One added bonus was the added finding of predictors of variance utilizing the Grading Utah Schools' sub-scores to predict school grades.

CHAPTER 5: CONCLUSIONS AND DISCUSSION

Purpose Statement

The purpose of this research was to determine the extent to which student demographics and school characteristics predict elements of a school accountability system. Specifically, this study applied school-level data to assess the level at which school characteristics such as low SES, English Language Learners, disability, racial minority, and school characteristics (such as the number of teachers with a graduate degree and rural settings) predict GUS grade, GUS overall points earned, GUS sub-scores, graduation rate, and College and Career Readiness ACT score.

Research Question

To what extent do student demographics and school characteristics predict school accountability grades?

Description of Research Design

Data used in this study for analyses included high school grades 9-11 tested in reading and math using the Student Assessment of Growth and Excellence (SAGE) during the 2014-2015 testing cycle. Of nearly 35,000 Utah students, 361 high schools were used for comparative and multiple regression analyses (Field, 2009; Stevens, 1999).

Comparative and multiple regression test analyses were used in computing bivariate correlations (Field, 2009; Stevens, 1999) in order to evaluate relationships between independent and dependent variables (George & Mallery, 2011). By using descriptive statistics (Stevens, 1999) and correlational analyses (Field, 2009), the research allowed for a view of the predictive power of school grades based on student backgrounds and school characteristics. After running multiple regression analyses

(Field, 2009; Stevens, 1999), the researcher determined the significant predictors of state accountability school grades.

Summary of Findings

Descriptive Statistics

Nearly 45% of teachers in participating schools held a graduate degree. The implications for having such a high percentage of teachers with a graduate degree poses additional questions for future research in investigating whether more affluent schools have more teachers with advanced degrees on staff. Almost 40% of students were identified as having a low SES. It would be critical to also research whether schools with higher populations of low SES students were negatively impacted with lower grades/rankings as a result of a lower percentage of teachers with graduate degrees serving in respective schools.

Racial minority students represented 15% of the total population. This suggests that the data is inconclusive in drawing conclusions on whether the racial minority group determines school grades (Davidson, 2016), but might be more significant in a different state with different demographics.

Multiple Regression Analysis

A standard multiple regression analysis (Field, 2009) was performed between school accountability grade and student and school characteristics (low SES, ELL, disability, racial minority, teachers with graduate degree, and rural school location). The regression analysis indicates that the independent variables, taken together, significantly predict GUS grades (Field, 2009). This supports the hypothesis that student and school characteristics can predict accountability results. This result is concerning, as little

attention is currently placed on these studies in education with regards to zoning of school districts and housing authority decisions, which can easily perpetuate the cycle of failure. When these characteristics are concentrated within a population, it pre-determines the achievement potential of a school. This connects to a problem noted in the literature: teachers transferring out of low-achieving schools into higher-achieving schools (Hanushek, Kain, & Rivkin, 2004).

The highest predictability for impacting school accountability grades connected both low SES and ELL students together along with teachers with a graduate degree. This is a concern because three of the independent variables were significant in determining the success of the school accountability grades. Knowing that schools with higher concentrations of poverty are less likely to have teachers with advanced degrees brings into question efforts to recruit and retain better teachers for higher need populations (Hundley, 2013). One important item was the statistical significance of teachers with a graduate degree and their impact on school grades. As a practitioner, there are professionals with advanced degrees who not always make gains with student achievement. However, this research supports the importance of hiring teachers with advanced degrees for optimal student achievement results (Darling-Hammond, 2000). Additional conversations are needed at the university level regarding rigorous preparation and intentional placements of student teachers within diverse settings, as well as thinking about incentives to attract more highly qualified teachers to the highest need schools. Further research would suggest a closer examination between school turnaround measures and an intentional focus on an equitable distribution of lower income students and ELL students across school districts. Also of importance, would be systematic

approaches in developing language acquisition to assist with high concentrations of ELL populations (Pinker & Prince, 1988), as well as ensuring more teachers with graduate degrees were strategically placed across diverse school settings, especially in working within low SES settings.

The researcher initially looked at how student and school characteristics predicted school grades. However, school grades are based on several different scores. Therefore, it is important to see if student and school characteristics predict those scores differently. Focusing on observing the scores from the tests that make up the student scores revealed whether certain parts of the test could be used to calculate school grades and determine if there were patterns of more or less bias against the various groups of students and types of schools. With that in mind, the researcher performed some additional analyses. A standard multiple regression analysis was also performed between Grading Utah Schools overall points earned and student and school characteristics. Analysis revealed that the model significantly predicted overall points earned, by 36% variance in overall points earned. Although this is moderate (Field, 2009), it is still significant. The three most significant predictors of success within the Grading Utah Schools points earned were ELL and students with disabilities along with the percent of teachers with a graduate degree.

For the English Language Arts Proficiency sub-score, the student characteristics percent low SES, percent English Language Learners, and percent disability significantly predicted points earned for Grading Utah Schools. With the obvious connection of the testing of English Language Arts negatively impacting ELL and students with a disability (Abedi, 2004; Johnson & Monroe, 2004) along with low SES backgrounds creating

disadvantages in learning (Michelson, 1972), it is worrisome to compare schools based on measured outcomes, when the student backgrounds are beyond the scope of the school's control (Toutkoushian & Curtis, 2005). With three demographic groups significantly predicting the outcome of an English Language Arts Proficiency sub-score, this proves the researcher's hypothesis correct and sounds an alarm for investing more time into more balanced assessments to represent students' learning in order to avoid continuing the cycle of failure and unfair judgments against schools with high concentrations of these populations.

A standard multiple regression analysis (Field, 2009) was then performed between graduation rate and student and school characteristics and revealed that the model significantly predicted overall points earned. Results indicated that 13% of the variance in graduation rate was accounted for by the variables. Although this is a small effect size (Field, 2009), it is still significant. ELLs and the percentage of teachers with a graduate degree each significantly predicted graduation rates. These findings support the researcher's hypothesis and review of literature in that ELLs have a better chance of success with regards to student achievement with repeated exposure to quality teachers with advanced degrees (Sanders & Rivers, 1996). These findings also confirm that student backgrounds are outside of the control of the school (Toutkoushian & Curtis, 2005) and a more concerted effort is in order for focusing on variables that can be controlled such as increasing teachers with a graduate degree numbers for students (Sanders & Rivers, 1996) and a more concentrated effort to ensure ELL students gain access to move through the graduation gateway, which directly impacts future income

(Greenstone, Harris, Li, Looney, & Patashnik, 2012). See Figure 1.1 for income expectation earned based on education.

In looking at College and Career Readiness ACT scores and student and school characteristics, the regression analysis (Field, 2009) revealed that the model significantly predicted overall points earned. Results indicated that 49% of the variance in school accountability grade was accounted for by the variables. This is a large effect size (Field, 2009) and significant. Low SES and students with a disability variables each significantly predicted ACT scores. These findings proved the researcher's hypothesis and calls into question the use of this assessment considering the predictability on school grades. Realizing the ACT is a gateway for student entrance to college and knowing the impact a degree has on potential earnings over the course of a student's lifetime, the researcher questions the fairness of this measure for low SES and students with a disability groups (Greenstone, Harris, Li, Looney, & Patashnik, 2012). See Figure 1.1 for income expectation earned based on education. This finding would make an argument that this testing practice could be perpetuating a social justice issue among students with low SES backgrounds and/or students with disabilities.

Implications

This research sought to determine if assigning grades to schools was equitable in ensuring an unbiased relationship to more than a ZIP code concerning student demographics or school characteristics. The researcher expected to find combinations of predictors of variance in determining school accountability grades, which were proven. However, the researcher did not expect such a strong predictor to be found in the importance for numbers of teachers with graduate degrees. Surprisingly, race and rural

locations were not predictors of performance in predicting school accountability grades. Overall, practitioners and policymakers must be aware that school demographics and characteristics are strong predictors of success of students with regards to education, future income, and determining the school accountability grades. With such a strong case in proving these predictors of variance, it is critical that more time is spent developing fair and just accountability models to be used in ascertaining the levels of growth and achievement for students. With the cycle of failure allowed to repeat for low SES, ELLs, and students with disabilities, it calls into question the reasons for allowing this kind of reporting to continue or why high concentrations of these groups are allowed to continue amassing within certain schools within school districts while other less diverse locations remain untouched with regards to high SES and homogenous populations. These findings call into question the political decisions behind the zoning of new properties; the opening of new schools; zoning of school districts; and local, state, and federal formulas that are used to determine the amount of money allotted to schools in addressing student needs across varying backgrounds.

Recommendations for Future Research

Educators are tasked with closing achievement gaps between various demographic groups that have diverse needs and challenges to overcome. Considering the literature review findings and results from the statistics run, it would be critical for school districts to look closely at how schools are districted along with taking a closer look at accountability models to ensure they are focused on continual growth and not just a benchmark performance. All schools are not equal in demographics, support, and teacher quality, so it would also be critical to spend additional research time looking into the

media's role on creating a story for the public, political decisions that further separate the *haves and have nots*, while also determining measured equitable supports for failing schools.

Concluding Remarks

The research confirmed the researcher's theory that student demographics and/or school characteristics were significant predictors of school accountability grades, as a whole as well as for the sub-score categories making up the points earned for grades. Student demographics that were negatively impacting student achievement that were significant were low SES, students with a disability, and ELL students. School characteristics that were most significant for positively impacting student achievement were the number of teachers with an advanced degree. The researcher's biggest surprise was the importance of seeing the impact of teachers with advanced degrees on student achievement and realizing that the literature showed that many of those teachers with advanced degrees transferred out of low-performing schools for high-performing schools (Greenberg & McCall, 1974; Murnane, 1981; Hanushek, Kain, & Rivkin, 2004).

This is one of the variables that can be controlled and improved in the interest of helping students and ultimately schools. Through the research, it would appear that the closing of achievement gaps is not merely a school or individual teacher issue, but rather an institutional and/or societal issue that cannot be adequately addressed until conversations about student groups, equitable funding, preparation of future educators, the way in which the media communicates information about schools, and political decisions all hone in on the real issues dealing with student achievement. By not

addressing the aforementioned, it would appear that current models are only perpetuating more social justice issues.

REFERENCES

Abedi, J. (2004). The No Child Left Behind Act and English language learners: Assessment and accountability issues. *Educational Researcher*, *33*, 4-14.

Acevedo-Garcia, D., Lochner, K., Osypuk, T., & Subramanian, S. (2003). Future directions in residential segregation and health research: A multilevel approach. *American Journal of Public Health*, *93*(2), 215-221.

Alcorta, J., Jr. (2011). The impact of community traits, school attributes and student demographics on student achievement. *Tarleton State University*, 1-150.

Armour, D. (2003). *Maximizing intelligence.* Somerset, NJ: Transaction.

Ballou, D. (2008, April). *Test scaling and value added measurement*. Paper presented at the National Conference on Value-Added Modeling, Madison, WI.

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, *28*(2), 117-148.

Bast, J., & Reitsma, P. (1998). Analyzing the development of individual differences in terms of Matthew effects in reading: Results from a Dutch longitudinal study. *Developmental Psychology*, *34*(6), 1373-1399.

Betebenner, D., & Linn, R. (2009, December). Growth in student achievement: Issues of measurement, longitudinal data analysis, and accountability. *Educational Testing Service*, 3-23.

Bogin, A., & Nguyen-Hoang, P. (2014). Property left behind: An unintended consequence of a No Child Left Behind "failing" school. *Journal of Regional Science*, *54*(5), 788-805. doi:10.1111/jors.12141

Booker, K., & Isenberg, E. (2008). *Measuring school effectiveness in Memphis.* Washington, DC: Mathematica Policy Research, Inc.

Bracey, G. (2004). The trouble with research, part 2. *Phi Delta Kappan*, *85*, 635-636.

Braun, H. (2008, September). *Vicissitudes of the validators.* Paper presented at the Reidy Interactive Lecture Series, Portsmouth, NH.

Briggs, D., & Betebenner, D. (2009). *Is growth in student achievement scale dependent?* Retrieved from http://www.cde.state.co.us/sites/default/files/documents/accountability/downloads/tap/growthscaledependent.pdf

Brown v. Board of Education, 347 U.S. 483 (1954).

Campbell, D. T. (2010). Assessing the impact of planned social change. *Journal of MultiDisciplinary Evaluation*, *7*(15), 3-43

Carr, P. J., & Kefalas, M. J. (2009). *Hollowing out the middle: The rural brain drain and what it means for America.* Boston: Beacon Press.

Chapin, J. R. (2006). The achievement gap in social studies and science starts early: Evidence from the early childhood longitudinal study. *The Social Studies*, *97*(6), 231-238.

Choi, K., Goldschmidt, P., & Yamashiro, K. (2006). *Exploring models of school performance: From theory to practice CSE 673.* Los Angeles, CA: University of Los Angeles/CRESST.

Colvin, R., & Helfand, D. (2000, July 1). Millions for schools tied to Stanford 9 test scores. *Los Angeles Times*. Retrieved from http://articles.latimes.com/2000/jul/01/news/mn-46763

Condie, S., Lefgren, L., & Sims, D. (2014). Teacher heterogeneity, value-added, and

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, *8*(1), 1-44.

Dorn, S., & Ydesen, C. (2014). Towards a comparative and international history of school testing and accountability. *Education Policy Analysis Archives*, *22*(115), 1-9.

Education Commission of the States. (1998). *Designing and implementing standard-based accountability systems.* Denver, CO: Author.

Education Week. (2016, January 4). *The Every Student Succeeds Act: Explained*. Retrieved from http://www.edweek.org/ew/articles/2015/12/07/the-every-student-succeeds-act-explained.html?tkn=UYZFzxES2YHsDD1WgTTRaV8EzSbF1msa9R%2FW&print=1

Egebeen, D. J., & Licter, D. T. (1991). Race, family structure, and changing poverty among American children. *American Sociological Review*, 801-817.

Elementary and Secondary Education Act (ESEA) of 1965, Pub. L. No. 89-10, §§ 201-604 Stat. 79 (1967).

Every Student Succeeds Act of 2015, 20 U.S.C. §§ 114-95 (2015).

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: SAGE Publications Ltd.

Figlio, D. N., & Getzler, L. S. (2002). *Accountability, ability and disability: Gaming the system.* Retrieved from http://www.nber.org/papers/w9307.pdf

Figlio, D. N., & Lucas, M. E. (2000). *What's in a grade? School report cards and house prices.* Retrieved from http://bear.warrington.ufl.edu/figlio/house0502.pdf

Ford, R., & Rice, K. (2015). Value-added results for public virtual schools in California. *Journal of Educational Technology & Society*, *18*(4), 412-423.

Fowler, W., & Walberg, H. (1991). School size, characteristics and outcomes. *Educational Evaluation and Policy Analysis*, *13*, 189-202.

Frankenberg, E., Lee, C., & Orfield, G. (2003). *A multiracial society with segregated schools: Are we losing the dream?* Retrieved from https://civilrightsproject.ucla.edu/research/k-12-education/integration-and-diversity/a-multiracial-society-with-segregated-schools-are-we-losing-the-dream/frankenberg-multiracial-society-losing-the-dream.pdf

Freel, A. (1998). Achievement in urban schools: What makes the difference? *Education Digest*, *64*(1), 17.

George, D., & Mallery, P. (2011). *SPSS for Windows step by step: A simple guide and reference, 18.0 update* (11th ed.). Boston, MA: Allyn and Bacon.

Gerald, D., & Haycock, K. (2006). *Engines of inequality: Diminishing equity in the nation's premier public universities*. Retrieved from http://edtrust.org/wp-content/uploads/2013/10/EnginesofInequality.pdf

Gibbs, R. (1995). Going away to college and wider urban job opportunities take highly educated youth away from rural areas. *Rural Development Perspectives*, 10, 35-44.

Glynn, T. R., & Waldeck, S. E. (2013). Penalizing diversity: How school rankings mislead the market. *Journal of Law & Education*, *42*(3), 417-500.

Goldschmidt, P., & Choi, K. (2007). *The practical benefits of growth models for accountability and the limitations under NCLB.* Retrieved from http://programs.ccsso.org/content/pdfs/P%20goldschmidt%20Policy%20Brief%20v3.pdf

Goldschmidt, P., Roschewski, P., Choi, K., Auty, W., Hebbler, S., Blank, R., & Williams, A. (2005, October). *Policymakers' guide to growth models for school accountability: How do accountability models differ?* Retrieved from http://www.ccsso.org/Documents/2005/Policymakers_Guide_To_Growth_2005.pdf

Greenberg, D., & McCall, J. (1974). Teacher mobility and allocation. *Journal of Human Resources*, *9*(4), 480-502.

Greenstone, M., Harris, M., Li, K., Looney, A., & Patashnik, J. (2012, September). *A dozen economnic facts about K-12 education.* Retrieved from http://www.hamiltonproject.org/assets/legacy/files/downloads_and_links/THP_12EdFacts_2.pdf

Hall, D. (1998). *Class size and demographics: What 3rd grade test results suggest about their impact on achievement in New Hampshire public schools.* Unpublished manuscript, The New Hampshire Center for Public Policy Studies.

Hanushek, E., Kain, J., & Rivkin, S. (2004). Why public schools lose teachers? *Journal of Human Resources*, *39*(2), 326-354.

Hanushek, E., & Rivkin, S. (2006). Teacher quality. *Handbook of the Economics of Education*, 2, 1051-1078.

Hanushek, E. A. (2004, April-May). The effect of school accountability systems on the level and distribution of student achievement. *Journal of the European Economic Association*, *2*(2-3), 406-415.

Hanushek, E. A., & Raymond, M. E. (2003a). *Improving educational quality: How best to evaluate our schools?* Retrieved from http://hanushek.stanford.edu/sites/default/files/publications/Hanushek%2BRaymond%202003%20Educ21stCent.pdf

Hanushek, E. A., & Raymond, M. E. (2003b). *Lessons about the design of state accountability systems*. Retrieved from http://hanushek.stanford.edu/sites/default/files/publications/Hanushek%2BRaymond%202003%20NCLB.pdf

Hanushek, E. A., & Raymond, M. E. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, *24*(2), 297-327.

Hundley, C. (2013, May). *Education and urban schools*. Retrieved from http://www.apa.org/pi/ses/resources/indicator/2013/05/urban-schools.aspx

Jacobsen, M. (2014, December 14). *Utah Office of Education releases 2014 school "grades".* Retrieved from https://www.ksl.com/?sid=32767383

Jaggia, S., & Tuerck, D. (2000). *Promoting good schools through wise spending.* Unpublished manuscript, Suffolk University, Boston, MA.

Jerald, C. (2001). *Dispelling the myth revisted: Preliminary finding from a nationwide analysis of "high-flying" schools*. Washington, DC: The Education Trust, Inc.

Johnson, E., & Monroe, B. (2004). Simplified language as an accommodation on math tests. *Assessment for Effective Intervention*, 29, 35-45.

Jordan, H., Mendro, R., & Weersinghe, D. (1997, July). *Teacher effects on longitudinal student achievement: A preliminary report on research on teacher effectiveness.* Paper presented at the National Evaluation Institute, Indianapolis, IN.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, *80*(4), 437-447.

Kodrzycki, Y. (2001). Migration of recent college graduates: evidence from the national longitudinal survey of youth. *New England Economic Review*, 13-34.

Koedel, C., & Betts, J. (2008). *Value-added to what? How a ceiling in the testing instrument influences value-added estimation*. Retrieved from http://www.nber.org/papers/w14778.pdf

Kress, S., Zechmann, S., & Schmitten, J. (2011). When performance matters: The past, present, and future of consequential accountability in public education. *Harvard Journal on Legislation*, 185-234.

Kulm, G. (2007). The achievement gap: Do we need a new strategy? *School Science and Mathematics*, *107*(7), 267-268.

Ladd, H. F., & Walsh, R. P. (2002). Implementing value-added measures of school effectiveness: getting the incentives right. *Economics of Education Review*, *21*(1), 1-17.

Lane, D. M. (n.d.). *Regression toward the mean*. Retrieved from http://onlinestatbook.com/2/regression/regression_toward_mean.html

Lane, S., Parke, C., & Stone, C. (1998). A framework for evaluating the consequences of assessment programs. *Educational Measurement*, *17*(2), 24-28.

Langenfeld, K., Thurlow, M., & Scott, D. (1997). *High stakes testing for students: Unanswered questions and implications for student with disabilities.* Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

LegiScan. (n.d.) Legislative research: TX SB6, 2015-2016, 84th legislature. Retrieved from https://legiscan.com/TX/research/SB6/2015

Licter, D. T., Cornwell, G. T., & Egebeen, D. J. (1993). Harvesting human capital: Family structure and education among rural youth. *Rural Sociology*, *58*(1), 53-75.

Linn, R. L. (2000). Assessments and accountability. *Educational Researcher*, *29*(2), 4-16.

Lissitz, R. W., Doran, H., Schafer, W. D., & Willhoft, J. (2006). Growth modeling, value added modeling and linking: An introduction. In R. W. Lissitz (Ed.), *Longitudinal and value added modeling of student performance* (pp. 1-46). Maple Grove, MN: JAM Press.

McLanahan, S. A. (1988). *The role of mother-only families in reproducing poverty*. Retrieved from http://www.ssc.wisc.edu/cde/cdewp/88-48.pdf

Merton, R. (1968). The Matthew effect in science. *Science*, *5*(1), 56-63.

Michelson, S. (1972). Equal school resource allocation. *Journal of Human Resources*, 7, 283-306.

Mintrop, H., & Sunderman, G. (2009). *Why high stakes accountability sounds good but doesn't work—and why we keep on doing it anyway.* Retrieved from https://civilrightsproject.ucla.edu/research/k-12-education/nclb-title-i/why-high-

stakes-accountability-sounds-good-but-doesn2018t-work2014-and-why-we-keep-on-doing-it-anyway/mintrop-high-stakes-doesnt-work-2009.pdf

Moon, T., Callahan, C., & Tomlinson, C. (2003). Effects of state testing programs on elementary schools with high concentrations of student poverty: Good news or bad news? *Current Issues in Education*, *6*(8), 1-22.

Morgan, P. L., Farkas, G., & Hibel, J. (2008). Matthew effects for whom? *Learning Disability Quarterly*, *31*, 187-198.

Murnane, R. (1981). Teacher mobility revisited. *Journal of Human Resources*, *16*(1), 3-19.

National Commission on Testing and Public Policy. (1990). *From gatekeeper to gateway: Transforming testing in America.* Chestnut Hill, MA: Author.

Nelson, R. (1999). *Closing or widening the gap of inequality? The intended and unintended consequences of educational accountability systems for students with disabilities*. Unpublished manuscript, University of Minnesota, Minneapolis, MN.

No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002).

Novak, J., & Fuller, B. (2003, December). *Penalizing diverse schools? Similar test scores, but different students, bring federal sanctions*. Retrieved from http://nepc.colorado.edu/files/EPRU-0312-48-RW.pdf *(*

Oakes, J. (1990). *Multiplying inequalities: The effects of race, social class, and tracking on opportunties to learn mathematics and science.* Santa Monica, CA: RAND.

O'Connor, M. (1989). Aspects of differential performance by minorities on standardized tests: Linguistic and sociocultural factors. In B. R. Gifford, *Test policy and test*

performance: Education, language and culture (pp. 129-181). Boston, MA: Kluwer.

Olszewski-Kubilius, P. S., Lee, M., Ngoi, M., & Ngoi, D. (2004). Addressing the achievement gap between minority and nonminority chldren by increasing access to gifted programs. *Journal for the Education of the Gifted*, *28*(2), 127-158.

Ornstein, A., & Levine, D. (1989). Social class, race and school achievement: Problems prospects. *Journal of Teacher Education*, *40*(5), 17-23.

Pallas, A. M., Entwisle, D. R., Alexander, K. L., & Stluka, M. F. (1994). Ability-group effects: Instructional social, or institutional? *Sociology of Education*, 67, 27-46.

Phillips, A. E. (2015). *Quantifying quality: The effects of score transformation method and school demographics on school rankings under the Elementary and Secondary Education Act*. Retrieved from https://conservancy.umn.edu/bitstream/handle/11299/175300/Phillips_umn_0130E_16209.pdf?sequence=1&isAllowed=y

Phillips, M. (2000). Understanding ethnic differences in academic achievment: Empirical lessons from natioal data. In D. W. Grissmer & J. M. Ross (Eds.), *Analytic issues in the assessment of student achievement* (pp. 103-132). Washington, DC: National Center for Education Statistics.

Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition, 28*(1-2), 73-193.

Pitoniak, M. J., & Royer, J. M. (2001). Testing accommodations for examinees with disabilities: A review of psychometric, legal, and social policy issues. *Review of Educational Research*, *71*, 53-104.

Raudenbush, S. W. (2004). *Schooling, statistics, and poverty: Can we measure school improvement?* Princeton, NJ: Educational Testing Service.

Ready, D. (2012). Associations between student achievement and student learning: Implications for value-added school accountability models. *Educational Policy*, 1-29. doi:10.1177/0895904811429289

Reese, W. (2013). *Testing wars in the public schools.* Cambridge, MA: Harvard University Press. doi:http://dx.doi.org/10.4159/harvard.9780674075672

Robelen, E. (2002). An ESEA primer. *Education Week on the Web*, *21*(16), 28-29.

Rothstein, J. (2008a). *Student sorting and bias in value added estimation: Selection on observables and unobservables.* Cambridge, MA: National Bureau of Economic Research.

Rothstein, J. (2008b). *Teacher quality in educational production: Tracking, decay, and student achievement.* Retrieved from http://www.nber.org/papers/w14442.pdf

Sahlberg, P. (2010). Educational change in Finland. In A. Hargraves, A. Lieberman, M. Fullan, & D. Hopkins (Eds.), *Second international handbook of educational change* (pp. 323-348). Dordrecht, NE: Springer Netherlands.

Sanders, W. L., & Rivers, J. C. (1996). *Cumulative and residual effects of teachers on future student academic achievement*. Retrieved from http://www.cgp.upenn.edu/pdf/Sanders_Rivers-TVASS_teacher%20effects.pdf

Seltzer, M. H., Frank, K. A., & Bryk, A. S. (1994). The metric matters: The sensitivity of conclusions about growth in student achievement to choice of metric. *Educational Evaluation & Policy Analysis*, 16, 41-49.

Sherman, J. & Sage, R. (2011). Sending off all your good treasures: Rural schools, brain-drain, and community survival in the wake of economic collapse. *Journal of Research in Rural Education*, *26*(11) 1-14.

Smith, R., & Yen, W. (2006). Educational testing service options for measuring growth. In R. W. Lissitz (Ed.), *Longitudinal and value added modeling of student performance* (pp. 82-94). Maple Grove, MN: JAM Press.

Solórzano, R. (2008). High stakes testing: Issues, implications and remedies for English language learners. *Review of Educational Research*, *78*(2), 260-329.

Stanovich, K. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*(4), 360-407.

Stecher, B. (2006, September 13). "No Child" leaves too much behind. *The Washington Post*. Retrieved from http://www.washingtonpost.com/wp-dyn/content/article/2006/09/11/AR2006091100581_pf.html

Stecher, B., & Hamilton, L. (2002). Test-based accountability: Making it work better. *Education Week on the Web*, *21*(23), 32, 48.

Stevens, J. (1999). *Intermediate statistics: A modern approach* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc..

Stevens, J., & Zvoch, K. (2006). Issues in the implementation of longitudinal growth models for student achievement. In R. W. Lissitz (Ed.), *Longitudinal and value added modeling of student performance* (pp. 170-209). Maple Grove, MN: JAM Press.

Stinebrickner, R., & Stinebrickner, T. (2003). Understanding educational outcomes of students from low-income families. *The Journal of Human Resources*, *38*(3), 591-617.

Summers, A., & Wolfe, B. (1976). Which school resources help learning? Efficiency and equity in Philadelphia public schools. *IRCE Bulletin*, *11*(3), 1-17.

Sutton, A., & Soderstrom, I. (1999). Predicting elementary and secondary school achievement with school-related and demographic factors. *The Journal of Educational Research*, *92*, 330-338.

Taylor, L. (2015). Excel National commends Senator Taylor, Texas Senate, for passage of A-F school grading. *Global Education Journal*, 2, 189-194.

Thum, Y. (2003). Measuring progress toward a goal: Estimating teacher productivity using a multivariate multilevel model for value-added analysis. *Sociological Methods and Research*, 32, 153-207. doi:10.1177/0049124103257073

Thurlow, M., & Johnson, D. (2000). High-stakes testing of students with disabilities. *Journal of Teacher Education*, *51*(4), 305-314.

Toutkoushian, R. K., & Curtis, T. (2005). Effects of socioeconomic factors on public high school outcomes and rankings. *The Journal of Educational Research*, *98*(5), 259-271.

Traylor, N. (2013). *The implications of accountability models weighing the same student data differently*. Retrieved from http://encompass.eku.edu/cgi/viewcontent.cgi?article=1141&context=etd

U.S. Department of Education. (2008, April). *Growth models: Ensuring grade-level proficiency for all students by 2014*. Retrieved from http://www2.ed.gov/admins/lead/account/growthmodel/proficiency.html

Utah State Board of Education. (2016). *Welcome to SAGE*. Retrieved from http://www.schools.utah.gov/sage/

Utah State Office of Education. (2015, October 1). *SAGE testing and data reporting*. Retrieved from http://www.schools.utah.gov/assessment/SAGE/SAGETestingDataReporting.aspx

Utah State Office of Education. (2015). *2015 Utah Accountability Technical Manual*. Retrieved from http://www.schools.utah.gov/assessment/Accountability/TechnicalManual.aspx

Utah State Office of Education Assessment and Accountability. (n.d.). *Preparing Utah's students for college and career: New standards, new tests, new scores*. Retrieved from http://schools.utah.gov/assessment/SAGE/StandardsTestScores.aspx

Wake, B. (2001, October 17). Anger over testing sweeps the U.S. *The Ottawa Citizen*p. B3..

Walberg, H. J., & Tsai, S.-L. (1983). Matthew effects in education. *American Educational Research Journal*, *20*(3), 359-373. Retrieved from https://www.psychologytoday.com/files/u81/Walberg___Tsai__1983_.pdf

Williams, A. (2011). A call for change: Narrowing the achievement gap between white and minority students. *The Clearing House*, 84, 65-71. doi:10.1080/00098655.2010.511308

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, 11, 57-67.

Ysseldyke, J., Nelson, J. R., Christenson, S., Johnson, D. R., Dennison, A., Triezenberg, H., & Hawes, M. (2004). What we know and need to know about the consequences of high-stakes testing for students with disabilities. *Exceptional Children*, *71*(1), 75-95.

Yu, F., Kennedy, E., Teddlie, C., & Crain, M. (2007). Identifying effective and ineffective schools for accountability purposes: A comparison of four generic types of accountability models. *Council of Chief State School Officers*, 1-53.

Zlatos, B. (1994). Don't test, don't tell: Is "academic red-shirting" skewing the way we rank our schools? *The American School Board Journal*, *191*(11), 24-28.

Zvoch, K., & Stevens, J. (2008). Measuring and evaluating school performance: An investigation of status and growth-based achievement indicators. *Evaluation Review*, *32*(6), 569-595. doi:10.1177/0193841X08320398

VITA FALL 2016

TARA GRIMES ISAACS

Office of Professional Learning & Support Services
Curriculum & Instruction
Jefferson County Public Schools
4425 Preston Highway
Louisville, KY 40213

Tel: (502) 485-3963
tara.isaacs@jefferson.kyschools.us

EDUCATION

Ed.D, student in Educational Leadership and Policy Studies, Eastern Kentucky University, Richmond, KY, December 2016.

Master of Education Administration, Eastern Kentucky University, Richmond, KY, May 2008

Master of Science, Curriculum and Instruction, University of Tennessee, Knoxville, TN, May 1995.

Bachelor of Arts, Sociology, Minor: Elementary Education, University of Tennessee, Knoxville, TN, May 1994.

COLLEGE TEACHING EXPERIENCE

Fall 2016 Project-Based Learning for principal certification class, Southern Indiana University. Guest Speaker.

PREVIOUS TEACHING/LEADERSHIP EXPERIENCE

2016- Present	District Professional Development Coordinator
2010-2015	Elementary School Principal
2008-2010	Highly Skilled Educator at Kentucky Department of Education
2007-2008	Elementary Assistant Principal/Professional Staff Assistant
2006-2007	District English Language Arts Specialist
1994-2006	Teacher in Grades: Kindergarten, $2^{nd}/3^{rd}$, 3^{rd}, 4^{th} and $4^{th}/5^{th}$

PROFESSIONAL DEVELOPMENT

Ed Leader 21, Fall 2016, Denver, CO
High-Tech High, Spring 2016, San Diego, CA
UK Next Gen Leadership Academy, Fall/Spring 2014/2015, Lexington, KY
Center for Creative Leadership, 2012/2013, Greensboro NC/Frankfort KY

PROFESSIONAL REFERENCES

Dr. Robert Biggin, Professor, College of Education, Eastern Kentucky University, Richmond, KY, robert.biggin@eku.edu

Dr. James Bliss, Professor, College of Education, Eastern Kentucky University, Richmond, KY james.bliss@eku.edu

Dr. Carmen Coleman, Professor, College of Education / Center for Innovation in Education, University of Kentucky, Lexington, KY, carmen.coleman@uky.edu

Dr. Charles Hausman, Professor, College of Education, Eastern Kentucky University, Richmond, KY, charles.hausman@eku.edu

Dr. Tom Shelton, Executive Director, Kentucky Association of School Superintendents, Lexington, KY, tom.shelton@kysupts.org