VITA

Keith Thomas Bush

EDUCATION

2015	Indiana State University, Terre Haute, Indiana Ph.D. in Educational Leadership
2008	Indiana University, Indianapolis, Indiana M.S. in Education
2004	Indiana University, Bloomington, Indiana B.S. in Education

PROFESSIONAL EXPERIENCE

2011-present	New Albany-Floyd County School Corporation, New Albany, Indiana Scribner Middle School, New Albany, Indiana Principal
2008-2011	Monroe County Community School Corporation, Bloomington, Indiana Tri-North Middle School, Bloomington, Indiana Assistant Principal – Athletic Director
2007-2008	MSD of Perry Township, Indianapolis, Indiana Perry Meridian 6th Grade Academy, Indianapolis, Indiana 6th Grade Math/Science Teacher
2004-2007	MSD of Perry Township, Indianapolis, Indiana Rosa Parks-Edison Elementary School, Indianapolis, Indiana 5th Grade Teacher

MEASURING SCHOOL ACCOUNTABILITY: PERCEPTIONS FROM TEACHERS, PRINCIPALS, AND SUPERINTENDENTS

A Dissertation

Presented to

The College of Graduate and Professional Studies

Department of Educational Leadership

Indiana State University

Terre Haute, Indiana

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Keith Thomas Bush

May 2015

Keywords: Accountability, perceptions of accountability, teachers, principals, superintendents

COMMITTEE MEMBERS

Committee Chair: Terry McDaniel, Ph.D.

Associate Professor of Educational Leadership

Indiana State University

Committee Member: Brad Balch, Ph.D.

Professor of Educational Leadership

Indiana State University

Committee Member: Todd Bess, Ph.D.

Director, Indiana Association of School Principals

ABSTRACT

The purpose of this study was to identify the multiple factors that are utilized by accountability models and to identify the perceptions of Indiana educators in regard to what these professional educators believe should be utilized to assess school accountability. No Child Left Behind and the Adequate Yearly Progress (AYP) accountability model frameworks have had an impact on all public schools in the United States. Under AYP, high stakes testing results are utilized to provide acknowledgement to schools that meet pre-established thresholds, as well as to provide consequences for schools that fail to meet these requirements. This study is significant because it measures the perceptions of educators in what variables they believe should be utilized to determine whether or not schools meet specific accountability thresholds. Potential reliability issues that are correlated to specific accountability models can be identified through the analysis of the composite score differences between the different accountability variables. Descriptive statistics, one-way ANOVAs, and a simple multiple regression were used to analyze and disseminate the data collected for this study. This study was conducted by administering an electronic survey to all public school K-12 teachers, principals, and superintendents throughout Indiana.

A 17-item accountability perception survey was created to quantitatively measure the opinions and perceptions of teachers, principals, and superintendents as to the variables used measure school accountability. A total of 746 educators submitted responses to the Perception of Variables Utilized to Effectively Measure School Accountability Survey. Statistical analysis of

the data included descriptive analysis for selected items as well as inferential analysis that included one-way ANOVAs and a simple multiple regression. The composite score was tabulated from the survey to determine if significant differences existed among dependent variables that included position type, locality, socioeconomic level, and school letter grade received from the Indiana Department of Education. By collecting the responses of teachers, principals, and superintendents, a composite score was determined as to the perceptions of the variables utilized to measure school accountability. A one-way ANOVA found a statistically significant difference among the composite scores when separated by position type. The teacher composite mean scores were significantly lower than the principal and superintendent composite scores. A one-way ANOVA found no statistically significant differences between locality (urban, suburban, rural), socio-economic level (0%–33%, 34%–67%, 68%–100%), or school letter grade received by the Indiana Department of Education (A, B, C, D or F) and the accountability variable composite scores. A multiple regression analysis determined that demographic factors (locality, number of years at current building, socioeconomic status, school size, Indiana A–F letter grade) do not serve as significant predictors of the composite scores for principals.

ACKNOWLEDGMENTS

I am very blessed to have so many inspirational people in my life. All of these people have been instrumental in my doctoral journey, as well as my life journey. Without their support, love, and guidance, I would not be the person I am today.

My dissertation chair, Dr. Terry McDaniel, has been pivotal in helping me navigate through this doctoral study. He has increased my intellectual capacity and allowed me to challenge my skill set throughout the past three years.

I would also like to thank Dr. Brad Balch and Dr. Todd Bess for being a part of my dissertation committee. I am extremely fortunate to have been surrounded by experts who share the same passion. I am forever grateful for their guidance, support, and expertise.

The professors within the Educational Leadership Department at Indiana State University have earned my upmost respect. Dr. Donlan and Dr. Gruenert helped me shape, or reshape, the way I look at this profession. Thanks to Dr. Coleman-Brown for guiding me through statistics.

Dr. Michael Langevin and Mrs. Judith Barnes deserve my sincerest gratitude for their work in preparing my dissertation. Your guidance and support are much appreciated.

I would also like to recognize Dr. Bruce Hibbard who encouraged me to further my education and supported me through this journey. I could not ask for a more supportive leader who recognized my potential and gave me a chance to be a part of New Albany-Floyd County Schools.

My New Albany cohort also deserves a huge thank you. I am blessed to have been

surrounded by such amazing educators. The many hours of discussion, presentations, studying, and sharing allowed me to grow as a leader and as a person. Specific recognition needs to be given to Dr. Rhonda Roos, Dr. Jessica Waters, and Dr. Steve Griffin. I cannot even begin to express how much I appreciate our friendship and camaraderie.

I would also like to recognize my family for their love and support. Thanks to my Mom and step-dad, Denise and Vic Combs, along with my mother- and father-in-law, Sharon and Tom Mutchler for their support through this process. My Mom has been instrumental in shaping me into the man and father I am today.

I have been unbelievably blessed to have two amazing children in my life. Kadin and Kinley, I hope that I have made you proud. Thank you for being supportive of me as I navigated this journey. Even at such a young age, the love each of you has for your mother and I is amazing. I love both of you with all of my heart. Remember to dream big and never give up!

Most importantly, I want to thank the love of my life, Andrea. Words cannot express the love and gratitude I have for you. You have been my rock, my heart, and my soul mate. Thank you for sacrificing so much to allow me to complete this journey. Thank you for loving me unconditionally! I hope that I am the husband and father you deserve. I cannot wait to see what the next chapter in our lives has in store for us.

TABLE OF CONTENTS

ABSTRACT iii

ACKNOWLEDGMENTS v

LIST OF TABLES xiii

LIST OF FIGURES xv

INTRODUCTION 1

Perspectives Regarding School Accountability 1

Statement of Purpose for the Study 3

Significance of the Study 5

Research Questions 6

Null Hypotheses 7

Limitations of the Study 8

Delimitations of the Study 8

Definition of Terms 8

Summary 9

LITERATURE REVIEW 11

A Historical Overview of School Accountability 12

The Equality of Educational Opportunity of 1966 Study 13

The National Assessment of Educational Progress 15

A Nation at Risk 17

Goals 2000 18

Improving America's School Act of 1994 (1994) 20

The No Child Left Behind Act of 2001 21

Criticisms of NCLB 26

Types of Accountability Models 28

Status Model 28

Improvement Model 29

Growth Model 31

Value-Added Model 33

State Specific Accountability Models 35

Colorado 35

Tennessee 38

Indiana 41

Variables That Impact Student Achievement 44

Content and Performance Standards 44

State Standards Assessments 44

Interpretation of Results 45

School Size 47

Socioeconomic Status 49

Race 51

Gender 51

The Principal's Impact on Schools 52

Summary 54

METHODOLOGY 56
Research Questions 57
Quantitative Research Design 58
Sample59
Survey Design 60
Content Validity 61
Survey Reliability 61
Data Collection and Procedures 62
Statistical Analysis 62
ANALYSES OF DATA 65
Descriptive Analyses 66
Whole Sample Demographic Data 66
Whole Sample Descriptive Data–Variable Belief 67
Demographic Statistics by Position Type–Teacher 69
Descriptive Statistics by Position Type–Teacher 70
Demographic Statistics by Position Type–Principal 72
Descriptive Statistics by Position Type–Principal 72
Demographic Statistics by Position Type–Superintendent 74
Descriptive Statistics by Position Type–Superintendent 75
Demographic Statistics by Indiana A–F Accountability Grade–A 77
Descriptive Statistics by Indiana A–F Accountability Grade–A 78
Demographic Statistics by Indiana A–F Accountability Grade–B 80
Descriptive Statistics by Indiana A–F Accountability Grade–B 81

Demographic Statistics by Indiana A–F Accountability Grade–C 83

Descriptive Statistics by Indiana A–F Accountability Grade–C 84

Demographic Statistics by Indiana A–F Accountability Grade–D/F 86

Descriptive Statistics by Indiana A–F Accountability Grade–D/F 87

Demographic Statistics by SES: 0%–33% Free/Reduced Lunch 89

Descriptive Statistics by SES: 0%–33% Free/Reduced Lunch..... 90

Demographic Statistics by SES: 34%–67% Free/Reduced Lunch 92

Descriptive Statistics by SES: 34%–67% Free/Reduced Lunch..... 93

Demographic Statistics by SES: 68%–100% Free/Reduced Lunch 95

Descriptive Statistics by SES: 68%–100% Free/Reduced Lunch..... 96

Demographic Statistics by School/District Locality–Urban..... 98

Descriptive Statistics by School/District Locality–Urban 99

Demographic Statistics by School/District Locality–Suburban 101

Descriptive Statistics by School/District Locality–Suburban..... 102

Demographic Statistics by School/District Locality–Rural..... 104

Descriptive Statistics by School/District Locality–Rural 105

Inferential Test Results 107

Null Hypothesis 1 108

Null Hypothesis 2 109

Null Hypothesis 3 110

Null Hypothesis 4 111

Null Hypothesis 5 112

Summary 115

SUMMARY OF FINDINGS, RESULTS,
IMPLICATIONS AND RECOMMENDATIONS 117
Summary of Findings 117
Results 118
Research Question 1 118
Research Question 2 119
Research Question 3 120
Research Question 4 121
Research Question 5 121
Research Question 6 122
Implications 123
Implication 1: Current Indiana Accountability Practice 123
Implication 2: Accountability Practices Outside Indiana 124
Implication 3: Potential Changes to the Indiana Accountability Model 125
Implication 4: Effect of Demographic Factors 125
Implication 5: Need for Continued Discussion 126
Implication 6: Indiana's Proposed Accountability Model 128
Areas for Further Research 129
Summary 131
REFERENCES 133
APPENDIX A: INFORMED CONSENT 151
APPENDIX B: PERCEPTIONS OF VARIABLES UTILIZED TO EFFECTIVELY
MEASURE SCHOOL ACCOUNTABILITY 153

APPENDIX C: PERMISSION TO USE ILLUSTRATIONS 158

LIST OF TABLES

Table 1. Percentage of Educators in High Poverty Schools Teaching Outside Their Subject Expertise 50

Table 2. Whole Sample Data for Variable Belief Statement 68

Table 3. Teacher Sample Data for Variable Belief Statement 71

Table 4. Principal Sample Data for Variable Belief Statement 73

Table 5. Superintendent Sample Data for Variable Belief Statement 76

Table 6. Indiana A–F Accountability Letter Grade Sample Data for Variable Belief Statement–A School 79

Table 7. Indiana A–F Accountability Letter Grade Sample Data for Variable Belief Statement–B School 82

Table 8. Indiana A–F Accountability Letter Grade Sample Data for Variable Belief Statement–C School 85

Table 9. Indiana A–F Accountability Letter Grade Sample Data for Variable Belief Statement–D/F School 88

Table 10. Variable Belief Statement by School SES: 0%–33% Free/Reduced Lunch 91

Table 11. Variable Belief Statement by School SES: 34%–67% Free/Reduced Lunch 94

Table 12. Variable Belief Statement by School SES: 68%–100% Free/Reduced Lunch 97

Table 13. Variable Belief Statements by School Locality–Urban 100

Table 14. Variable Belief Statements by School Locality–Suburban 103

Table 15. Variable Belief Statements by School Locality–Rural 106

Table 16. ANOVA for the Regression Model for Demographic Factors Predicting the Composite Score of Principals 114

Table 17 Comparison of Mean Composite Scores Between Position Type with Significant Results 120

LIST OF FIGURES

Figure 1. State differences in AYP indicator interpretation. 25

Figure 2. Status model.. 28

Figure 3. Improvement model..... 30

Figure 4. Growth model.. 31

Figure 5. Value-added growth model.. 34

CHAPTER 1

INTRODUCTION

Perspectives Regarding School Accountability

Three days after taking the office of President of the United States, George W. Bush introduced his administration's agenda for education reform known as No Child Left Behind (NCLB) that was described as *the cornerstone* of the Bush administration (United States Department of Education [USDOE], 2002). The passage of NCLB (2002) marked an instrumental change in educational policy at the federal level. This change put an emphasis on accountability rather than an emphasis on standards and assessment. Standards and assessment are a critical component in NCLB; however, these items were used as a catalyst for schools, districts, and states to focus on every student's academic achievement. The goal of NCLB was to "ensure that all students showed academic proficiency on state's academic achievement standards and state academic assessments by 2013-14" (NCLB, 2002, p. 2).

President Bush stressed his profound confidence in U.S. public schools but shared an even deeper concern that "far too many of our neediest children are being left behind" (USDOE, 2002, p. 1). This was despite the fact that since the 1965 passage of the Elementary and Secondary Education Act (ESEA), nearly $200 billion in federal spending has gone toward public schools. Contrary to prior ESEA reauthorizations, the provisions set forth by NCLB were to become effective immediately after the enactment of the NCLB in January 2002 (Marion et

al., 2002). This provision required states to implement the majority of the changes outlined in NCLB by the beginning of the current school year.

The intention of the NCLB Act (2002) was to further support the Title I accountability framework by requiring states to implement statewide accountability systems that would oversee all public schools and students. Under NCLB all students, by 2013-2014, are expected to attain specific proficiency levels in mathematics and reading and language arts. All students can be defined as students, regardless of race, ethnicity, socioeconomic status, gender, disability, or English proficiency (CTB/McGraw-Hill, 2002; National Association of School Boards of Education, 2002).

As a means of measuring if each student meets academic proficiency, NCLB requires both schools and districts to make adequate yearly progress (AYP). NCLB quantifies AYP by utilizing annual measurable objectives (AMO). The AMO can be described as the proportion of students in a specific group that meet or exceed the proficient level on a standardized assessment. Per NCLB, an educational institution can be only be classified as to meeting AYP if the entire student population and each of the six subgroups, mentioned above, are able to meet the required AMO. As a means of accountability, academic institutions that are unable to meet the AYP requirements for two or more consecutive years will potentially face consequences (CTB/McGraw-Hill, 2002; National Association of School Boards of Education, 2002).

For the past 13 years, NCLB has been a critical topic in the educational profession. Proponents of NCLB argue that this relentless insistence of 100% proficiency gives educators, parents, and students a common goal (Imms, 2008). It is also argued that educational institutions that are failing students should be held accountable (Imms, 2008). In contrast, opponents of NCLB claim that this law actually leaves students even farther behind. The majority of the

criticism to NCLB can be related to the methods utilized in the determining of AYP (Doran & Izumi, 2004).

Statement of Purpose for the Study

NCLB was designed around one critical issue in education—accountability. The NCLB reform is a set of accountability standards for educational institutions set forth by the USDOE in response to the concern that students were not achieving. This study examined a portion of the NCLB accountability system, along with alternative accountability models, in an attempt to identify which variables members of the education profession believe should be utilized. Under NCLB, schools are held accountable through a measurement of student achievement levels in both the school and school district. The philosophical focus on every student meeting ambitious academic goals to close the achievement gap under NCLB has been commended (Borowski & Sneed, 2006; Chubb, 2005; Guilfoyle, 2006; Haycock, 2006; Hess, 2006; Hess & Petrilli, 2006; Lewis, 2006; Staiger, Geppert, & Kane, 2002). According to Guskey (2007), educational leaders should embrace accountability that focuses on student achievement to improve the quality of education in school, as a positive aspect of NCLB.

As a means of measuring that every student is learning, NCLB utilized AYP. AYP is an example of a status based accountability system that compares a subgroup's level or school's level of student proficiency at one point in time to an existing target level. AYP was widely accepted because it provided the public with an easy to understand snapshot of a subgroup or school's academic performance at a given time. However, criticisms of the AYP model led to further exploration of what data should be collected, as well as additional models that exist that could better assess student achievement. This study analyzed the effectiveness of the AYP accountability measurement system by comparing this system to alternative accountability

models and its variables used.

In response to criticisms toward AYP, along with others, the USDOE granted states a level of flexibility to NCLB. Most relevant to this study, is the flexibility that focuses on the use of different accountability models rather than the sole focus on the AYP status model. This study analyzed and compared different accountability models and the variables used, utilizing survey results from Indiana teachers, principals, and superintendents, to ascertain if there is a relationship in the belief of what variables should be used. The sample size also looked for a relationship in common responses grouped by socioeconomic status (SES), position type, location type, and each school/district's accountability grade.

This study evaluated educators' perceptions in regard to multiple accountability variables that could be utilized to deem a school effective or ineffective. This was measured using quantitative data analyses. Recent research has suggested that reliability issues exist due to the lack of consistency with utilizing AYP to measure school accountability (Elmore, 2002; Guilfoyle, 2006; Linn & Haug, 2002; Maleyko & Gawlik, 2011; Porter, Linn, & Trimble, 2005; Schoen & Fusarelli, 2008; Staiger et al., 2002; Wiley, Mathis, & Garcia, 2005). Linn (2008) argued that a state's ability to identify a successful school or a failing school may be inconsistent due to the methodology that is utilized by the state to measure AYP rather than the actual effectiveness of schools. This study focused on the variables that are used to measure school accountability and their significance as indicators of school performance. The NCLB Act of 2001 requires states to establish an accountability system that is both valid and reliable (NCLB, 2002). As evidenced by ongoing policy debates, the question of what, who, and how to measure school performance continues to plague the validity of NCLB (Gong, Perie, & Dunn, 2006).

Significance of the Study

NCLB (2002) and the AYP accountability model framework have an impact on all public schools in the United States. Under AYP, high stakes testing results are utilized to provide acknowledgement to schools that meet pre-established thresholds, as well as to provide consequences for schools that fail to meet these requirements. This study is significant because it measures the perceptions of educators in what variables they believe should be utilized to determine whether or not schools meet specific accountability thresholds. Potential reliability issues that are correlated to specific accountability models can be identified through the analysis of the composite score differences between the different accountability variables. The findings of this study could have significant implications in regard to the choice of specific accountability variables by states, specifically Indiana. Identifying schools that are ineffective and finding ways to assist schools in improving is at the heart of accountability. The use of different accountability variables has the potential to define school effectiveness in different ways. Therefore, schools can be defined effective in one specific model and be held to a completely different set of standards in another model, resulting in being deemed ineffective. By highlighting the best procedures for determining effective/ineffective schools and identifying variables that have consistent outcomes for all Indiana schools will be of methodological importance.

With the increased utilization of student performance as a means of holding educational institutions accountable, as witnessed within NCLB (2002), states should have a reliable accountability system that can accurately and fairly measure schools. Included in this study is scientific data regarding the reliability of using specific accountability variables. This study compared different accountability variables, as well as the beliefs of Indiana educators on which

variables should be utilized and the findings should have specific implications regarding educational accountability policy choices for Indiana legislators. The impact of NCLB legislation continues to grow, and the continued reauthorization of specific provisions within NCLB proves that it will continue to be a part of accountability within education for the future. This information could be useful in determining if NCLB reform, and its flexibility, is meeting the goals that were set out by the legislature at both the federal and state levels.

Research Questions

According to Creswell (2003), "Quantitative research questions inquire about the relationships among variables that the investigator seeks to know" (p. 132). The study consisted of six research questions. Each question addressed a specific grouping of survey responses. This study utilized the data collected from a quantitative survey sent to all Indiana teachers, principals and superintendents. Question 1 looked at the current perceptions of Indiana educators of the variables used in various accountability models. Question 2 looked at the differences in composite scores based on of the position type of each respondent. Question 3 addressed the differences in composite scores and the school or district accountability grade. Question 4 utilized a school or district's socioeconomic status to observe differences in survey results. Question 5 compared the results of composite scores for survey respondents based on location type. Finally, Question 6 addressed the ability to predict the accountability composite scores utilizing five different variables.

1. What are the current perceptions of Indiana educators on the variables in various accountability systems?
2. Is there a significant difference in the accountability composite score based on the type of position of the respondents?

3. Is there a significant difference in the accountability composite score based on the school or district's accountability grade?
4. Is there a significant difference in the accountability composite score based on the school or district's socioeconomic level?
5. Is there a significant difference in the accountability composite score based on the school or district's location type?
6. Does school socioeconomic status level, school or enrollment size, years in current position as a principal, school location, and school accountability grade serve as predictors of the composite score?

Null Hypotheses

$H_0$1. There is no significant difference on the accountability composite score based on educational position type.

$H_0$2. There is no significant difference on the perception composite score based on accountability letter grade received.

$H_0$3. There is no significant difference on the perception composite score based on SES level.

$H_0$4. There is no significant difference on the perception composite score based on the school or district's location type.

$H_0$5. The socioeconomic status level, school enrollment size, years in current position as principal, location, and current school accountability grade does not serve as predictors of the composite score.

Limitations of the Study

1. The Indiana teacher, principal, and superintendent e-mail database was gathered from the Indiana Department of Education (IDOE). The accuracy of this database could not be assured.
2. The accuracy of the respondents could not be controlled.

Delimitations of the Study

1. Only K-12 teachers, principals, and superintendents employed during the 2014-15 school year in Indiana were surveyed.
2. The enrollment in their schools, specific socioeconomic levels, and the number of years in their current position were the only controls for principals.
3. The school accountability letter grade utilized data from the 2012-2013 school year.

Definition of Terms

End of course assessment (ECA): The ISTEP+, criterion-referenced assessments developed specifically for students completing their instruction in Algebra I, Biology I, or English 10.

Growth models: The growth model is a type of accountability model utilized to measure student growth over a period of time. Growth is measured by observing the change in score from one year to the other. The actual growth is then compared to a prescribed expected growth (Goldschmidt et al., 2005; Yu, 2007).

Improvement models: An improvement model compares the change in student or district proficiency at two points in time (Goldschmidt et al., 2005; Riddle, 2005; Yu, 2007). The difference between the proficiency levels provides a measure of improvement.

Principal: A principal refers to a person who has direct authority over a school consistent

with district policy, goals, and objectives.

Safe Harbor: The secondary indicator of school performance under NCLB. According to Goldschmidt et al. (2005), Safe Harbor is utilized if a school fails to make AYP for a given school year, yet reduces the number of students scoring below proficiency by 10% from the previous year's comparable group (Riddle, 2005).

School accountability models: Accountability models target either the procedures or outcomes of education. Specific goals are identified and procedures are identified for measuring whether the goal is met.

School accountability variables: The specific criteria used for determining whether the goal has been met. These variables "can include specific determinations of ways that the goal may and may not be met" (Thurlow, 2009, p. 1).

Status models: The status accountability model utilizes student achievement data at one point in time, and compares that data with a pre-established target (Goldschmidt et al., 2005; Yu, 2007).

Superintendent: The superintendent is person who is the leader of a school district or corporation.

Teacher: A teacher is a person whose responsibility is to teach students.

Value-added models: Similar to growth models, value-added attempt to measure both teaching and learning. By utilizing multiple variables and looking at growth throughout the history of the student, connections are made to the impact a teacher or school had on a particular student (Lissitz, Doran, Schafer, & Willhoft, 2006).

Summary

Focusing on educators in the state of Indiana, this study utilized a quantitative design to

measure the current perceptions on the variables utilized in various accountability systems. In an attempt to identify relationships, the perceptions were disaggregated by additional response variables: educator position type, SES level, location type, and the 2012 – 2013 Indiana A – F Accountability designation. The purpose of this study was to identify the multiple factors that are utilized by these accountability models and to identify the perceptions of Indiana educators in regard to what they believe should be utilized to assess school accountability.

Chapter 2 of this study provides a historical examination of the accountability movement throughout the United States. This chapter also examines the evolution of accountability models from the beginning of NCLB (2002) to present. Additionally, this chapter examines the historical perspective of three specific geographic locations (Indiana, Colorado, and Tennessee and their transformations to their accountability systems. Finally, Chapter 2 examines common variables utilized in different accountability models and the impact of these variables on the accountability outcomes. The research that is referenced includes empirical studies and theoretical references from researchers who addressed the implementation of accountability models by, both, the federal and state governments. A thorough examination of the research was conducted to analyze the perspectives that support the design of this study.

Chapter 3 provides a detailed description of the research methodology. This methodology consists of a quantitative analysis of survey results gathered from Indiana educators. The quantitative analysis provides answers to the each of the research questions. Chapters 4 and 5 reveal the findings of this study and recommend future research studies as it relates to measuring school accountability.

CHAPTER 2

LITERATURE REVIEW

Divided into four main sections, this chapter focuses on (a) the historical progress of school accountability, (b) NCLB legislation, (c) state-specific school accountability systems, and (d) studies that compare variables that are commonly used in accountability models. Section 1 looks at the history of the accountability reform movement. School accountability began as an attempt to measure and monitor the performance of individual schools and then morphed into school accountability when the results of standardized testing were attached. The transformation of school accountability has taken place due to several major events. These include the "publication of the Coleman Report (1966), the creation of The National Assessment of Education Progress (NAEP), the publication of A Nation at Risk" (Yu, 2007, p. 14), and the creation and multiple renewals of the ESEA (1965). Also examined are the positives as well as the negatives of attaching standardized testing to accountability.

Section 2 is an examination of the NCLB Act of 2001 (NCLB; 2002). This section identifies accountability provisions found in NCLB as well as acknowledges criticisms of NCLB. Finally, this section also examines various flexibility options granted under NCLB.

Section 3 addresses the shift in accountability systems in the state of Indiana, Tennessee, and Colorado. The state of Indiana was selected due to its use of the Indiana A – F growth model. The state of Colorado has gone through an extensive shift in their accountability

framework. Finally, the state of Tennessee has been utilizing a value-added model for a number of years in which multiple variables are observed.

Section 4 examines common variables utilized in different accountability models and the impact of these variables on the accountability outcomes. The overall findings among this research are that the utilization of simple variables, like yearly snapshots of student performance, provides distorted information about the overall effectiveness educational institutions. The utilization of multiple variables, such as a growth model or value-added model, provides an alternative option to measure school accountability.

A Historical Overview of School Accountability

Tracing the origin of school accountability can be difficult. Basic methods of addressing school accountability began to appear as early as the 19th century with the creation of tax-supported public schools. According to Cuban (2004), community members were elected to serve as trustees on local school boards. In this position, members were duty-bound to "ensure that public school children were adequately housed, taught and supplied with learning materials" (Yu, 2007, p. 15). More of the focus of school accountability was directed on how money was spent, rather than student learning. The quality of schools was defined as "one that efficiently provided students access to adequate buildings, staff and materials" (Cuban, 2004, p. 28).

Equitable access to resources was the main focus of initial educational reform. The general belief held that more supplies, higher quality textbooks, newer facilities, and higher teacher salaries would assist in fixing the gaps within America's schools. However, three high profile events led to the focus on school accountability. The release of Coleman's (1966) *Equality of Educational Opportunity* study, the creation and implementation of the NAEP, and

the publication of *A Nation at Risk* (National Commission on Excellence in Education, 1983) which brought school accountability to the forefront of public education.

The Equality of Educational Opportunity of 1966 Study

Interest in student achievement-based accountability can be traced back to the Equality of Educational Opportunity of 1966 report by author and sociologist James Coleman. Authorized by the U.S. Department of Health, Education, and Welfare in 1966, The Equality of Educational Opportunity Study (EEOS) was conducted to assess the availability of equal educational resources to students of all races, religions, colors, and national origins. The EEOS study is often referred to as the Coleman Study. Conducted in response to requirements set forth in the Civil Rights Act (1964), the Coleman Study served as one of the first examples of educational policy being made around the results of data instrument. The EEOS analyzed test scores and survey responses obtained from first-, third-, sixth-, ninth-, and 12th-grade students, and survey responses from educators. Utilizing a national sample of schools in the United States, data included student age, gender, race and ethnic identity, socioeconomic background, attitudes toward learning, education and career goals, and racial attitudes (Coleman, 1966). Also included in this report were performance scores from various standardized tests that assessed verbal skills, nonverbal associations, reading comprehension, and mathematics (Coleman, 1966).

In this report, Coleman (1966) compared the distribution and opportunities among children of different races. Also located in this report was an in-depth examination of outcomes in student achievement scores in regard to race. Contrary to perceptions at the time, the Coleman Study found that the home environment and family background had a larger impact on student achievement than the provided educational resources (Coleman, 1966). This report marked a shift in education, which had previously looked at equity and accountability at the resource, or

input level, to a concentration on student learning, or output. This led research to focus on the effect that school resources had on student achievement (Ravitch, 1983). The Coleman Study led to national conversations centered around the inequalities students came to school with and ways to overcome these disparities (Coleman, 1966).

The response to the Coleman report brought forth new views in regard to education and social class. As Marshall (1998) remarked,

> Coleman himself later argued that the most important research findings of the study were twofold. First, it showed that variations in school quality (as indexed by the usual measures such as per pupil expenditure, size of school library, and so on) showed little association with levels of educational attainment, when students of comparable social backgrounds were compared across schools. (Differences in students' family backgrounds, by comparison, showed a substantial association with achievement.) Second, a student's educational attainment was not only related to his or her own family background, but also (less strongly) to the backgrounds of the other students in the school. (in intro to quote, para. 2)

After the Coleman report, student achievement and proficiency, along with resources, became the focus of educational reformers. According to Ravitch, "reformers advanced a broader array of proposals, many which sought changes in performance rather than increases in resources" (as cited in Evers & Walberg, 2002, p. 15). Policymakers and school and community officials began experimenting with different structural changes of the education system that were intended to have an impact on student performance, or outcome. These changes included the decentralization of schools, offering school choice to parents, as well as contracting with private companies to deliver services to students of low-income areas. Again, another first in education

occurred, in that the whole concept of education was being questioned and the structural design was challenged. These challenges were exacerbated by a call for more in-depth reviews of student achievement data. This request for additional data led to Congress authorizing NAEP.

The National Assessment of Educational Progress

The role of the federal government in education increased dramatically during the 1960s. Following the assassination of John F. Kennedy in 1963, Lyndon Johnson won a landslide victory in the presidential election on the platform of the Great Society. With his vow to fight a war on poverty, President Johnson was able to pass the ESEA of 1965. ESEA helped pave the way to allow the federal government to intervene in what was once a state dominated education system. During the 1960s, concerns about ways to assess students reflected the growing interest in accountability with the federal government. As a provision of ESEA, the federal government began planning and developing a national system for assessing student ability and knowledge.

A former dean of the Harvard School of Education, Francis Keppel played an instrumental role in the creation of NAEP test. As the U.S. Commissioner of Education, Keppel was frustrated with the department's inability to identify the effects of resources on student achievement. As Keppel (1966) remarked,

> It became clear that American education had not yet faced up to the question of how to determine the quality of academic performance in the schools. There was a lack of information. Without a reporting system that alerted state or federal authorities to the need for support to shore up educational weakness, programs had to be devised on the basis of social and economic data. . . . Economic reports existed on family needs, but no data existed to supply similar facts on the quality and condition of what children learned. The nation could find out about school buildings or discover how many years children

> stay in school; it had no satisfactory way of assessing whether the time spent in school was effective. (pp. 108-109)

In the spring of 1969, the first NAEP data collection was a trial assessment of the citizenship, science, and writing performance of 17-year-old in-school students. During the fall of that same year, assessment data of 9- and 13-year-old students as well as out-of-school 17-year-olds were collected. The NAEP framework described the content and skills measured on NAEP assessments as well as the design of the assessment. They provided both the *what* and the *how* for national assessment. The NAEP assessment is, currently, the only national assessment that is used to measure what students in the United States know and can do in specific subject areas. However, as Peterson and West (2003) stated, "But even though it was designed so as not to inform anyone about how individual schools were doing, NAEP, ironically, would prove to be a key mechanism in hastening the accountability movement forward" (p. 5).

The results of NAEP, including information on student performance, are published as The Nation's Report Card. These data are often broken down into specific subcategories, such as age, race, socioeconomic status, and demographics. As part of NCLB (2002) provisions, an annual assessment of content standards in reading, mathematics, and science is required to be administered by each specific state. NCLB also requires that students in Grades 4 and 8 must take the NAEP assessment in reading and mathematics every other year if their specific state receives Title I funding. Educational institutions that administer the NAEP assessment do so in same way nationwide. By following the same protocol and asking the same questions across the nation, NAEP provides a common measurement of student progress and makes it possible for comparisons of states data to be made. The results attained from the NAEP assessments, along with individual state assessments, give entities the ability to monitor academic progress and

develop ways to improve education in America.

A Nation at Risk

After analyzing data from the NAEP examinations, a federal report by the National Commission on Excellence in Education was issued. This report, entitled *A Nation at Risk*, was published in April 1983. *A Nation at Risk* asserted that American "students were not studying the right subjects, were not working hard enough, and were not learning enough. Their schools suffered from slack and uneven standards. Many of their teachers were ill-prepared" (Finn, 1989, p. 17). *A Nation at Risk* cautioned that "our social structure would crack, our culture erode, our economy totter, [and] our national defenses weaken" (Finn, 1989, p. 17) if work to change the structure of the current education system in the United States did not immediately take place.

Called "the most important education reform document of the 20th century" by education historian Diane Ravitch (Toppo, 2004), A Nation at Risk identified many areas that raised concerns about the American education system and predicted that America would soon be consumed by a "rising tide of mediocrity in elementary and secondary school" (Kantrowitz, 1993, p. 46). A complete overhaul of public education was called for. *A Nation at Risk* has catapulted education into an enduring topic on the national stage, rather than just a local and state issue. Every president, since this publication, has included educational reform in his political agenda. "The business community, weary of having to run its own remedial programs, [has also] joined the crusade for a better-trained work force to compete in a global economy" (Kantrowitz, 1993, p. 46). With education in the public spotlight, many states and local governing bodies have pushed for comprehensive education-reform. This legislation adds to graduation requirements, require teachers to take content area exams, redesign teacher licensing

requirements, decrease the average class size, and much more (Finn, 1989). Along with these changes, calls for the increase in average student spending, as well as the belief that raising teacher salaries would have an effect on student achievement were also results brought on by this report. A Nation at Risk served as an eye-opening spark to the national education reform debate.

Goals 2000

In 1992, Bill Clinton was elected president. President Clinton was a popular Democratic governor of Arkansas who participated in the 1989 education summit and was an advocate for the creation of educational goals at the national level. Clinton's first legislative proposal was called *Goals 2000: The Educate America Act* (Educate America Act, 1993). Goals 2000 was introduced in 1993 and believed to have been influenced by the National Council on Education Standards and Testing, the Goals Panel, and the experience of states already implementing educational reform (Roe, 2009). Aligning with the National Education Goals Panel's six national standards, Congress approved legislation along with two additional goals, which focused on teacher quality and parental responsibility (Educate America Act, 1993). According to Roe (2006), early advocates of the national goals wanted the focus of Goals 2000 to be focused solely on student performance. Thus, the addition of teacher quality and parental responsibility was met with opposition. In February 1994, and a year into President Clinton's first term, The Goals 2000 legislation passed.

The Goals 2000 framework was considered a grant program utilized by states to support the development of content specific standards and assessments. Also located in the framework was financial support to assist school districts with implementing standards-based reform. Legislation did not consider Goals 2000 a direct federal program, thus sparse regulation existed to oversee its implementation. Many states were already working toward systematic reform and

Goals 2000 was an attempt to show support for these efforts. Under Goals 2000, any state that was attempting to work toward the concept of a standards-based, systemic reform movement and had a planning process in place to support this movement could receive funding. Unlike previous federal programs that targeted specific subject areas or groups of students, Goals 2000 was unusual in that it supported reform strategies that focused on the development of state standards and the assessments needed to measure progress toward them (Roe, 2009). It required that most of the funds go toward the implementation of state standards at the local school or district level.

The reliance of Goals 2000 on state-by-state initiatives meant that its impact varied greatly between states, between districts and even between schools (Superfine, 2005). When considering the fact that the majority of financial support had to be spent at the district level along with the sheer number of districts, funding was not readily available to move resources to districts that were already struggling. Furthermore, difficulties in creating support for teachers to develop, implement and utilize the new standards began to occur. This, again, was due in part to the nearly 5,000 districts distributing a relatively small amount of funds. In 1997, the Urban Institute initiated a study (Hannaway, 1997) that analyzed the use of Goals 2000 grant money. Their study found that schools who struggled with the implementation of standards-based reform were typically small and/or high-poverty districts (Hannaway, 1997). Specifically, districts reported struggling with understanding the standards-based education reform process and lacked the resources necessary to seek out support. It is also noted that high-poverty districts were informed and aware of the federal program; however, these districts felt they did not have the expertise to provide technical assistance to schools or connect to essential information on this topic (Hannaway, 1997).

Small districts also struggled to adapt to the standards-based reform movement. Similar to schools servicing high numbers of students from poverty, data suggests that small schools lacked the expertise and information needed to adjust to the standards-based reform movement. Even though some local districts experienced difficulties, Goals 2000 was instrumental in providing the funding to support the development and implementation of the standards-based reform movement. State governors and business leaders worked with local education agencies to develop these educational reforms. Initially, the funding for Goals 2000 was set at $94 million in 1994. However, with staunch support from education and business groups, funding soared to $490 million by 1999. Overall, the standards-based reform movement received close to $2 billion in funding through Goals 2000. Historically, education funding had been primarily funded through each individual state. With this in mind, some states initially turned down federal funding through Goals 2000. However, within the first two years of the legislation passing, 48 states accepted federal grants. Within three years, more than 5,000 school districts across the country accepted Goals 2000 funding to address standards-based reform.

Improving America's School Act of 1994 (1994)

In conjunction with Goals 2000, President Clinton moved forward with proposals for the reauthorization of ESEA, now called Improving America's Schools Act or IASA (1965). Prior to modifying ESEA, the President wanted to setup the framework for new state standards and test. Fearful that any additional funding that was acquired for Goals 2000 would be used to support only Title I (under ESEA), it was critical to President Clinton that Goals 2000 and IASA not be combined together. The intent of both bills was a much broader systemic reform that would support all schools (Roe, 2006). President Clinton triumphed. Goals 2000 was passed, and in October 1994, IASA was signed into law.

Under previous authorizations of ESEA (1965), Title I provisions allowed states to use different academic standards for economically disadvantaged students than the ones used for regular students. These standards were often less rigorous than the normal standards. With the passage of IASA, all students, both Title I and non-Title I, were required to meet the new standards being created through Goals 2000. The combination of both pieces of legislature helped states move forward with standard-based reform for all public-school students by directing close to $11 billion into federal grants (Roe, 2006).

A clear focus on the standards of learning led to IASA to receive support from both Republicans and Democrats in a bipartisan fashion. According to Roe (2006),

> The education and business communities and the Commission on Chapter 1 (currently known as Title I), whose report described the Chapter 1 program as focusing mostly on low-level programs with little accountability for improved outcomes among disadvantaged students, were extremely supportive of this new law. (p. 67)

With the passage of IASA, *Title 1* (formerly Chapter 1) started servicing low-achieving and low-income students as one of the largest federal grants program ever. This outcome was viewed as one of the most significant utilizations of federal power in educational policy; power that had previously been held at the local and state level (P.L. 103-382). IASA and Goals 2000 became "the largest single federal funding stream for elementary and secondary education, depending on state and local decisions around standards, testing, teacher training, curriculum, and accountability" (Roe, 2006, p. 67).

The No Child Left Behind Act of 2001

On January 8, 2002, President Bush signed his reauthorization of ESEA, NCLB (2002), into law. All 50 states quickly became faced with an unprecedented challenge. NCLB called for

the implementation of a prescribed accountability model with the goal that all students would achieve grade-level proficiency within 12 years. Traditionally, education had been controlled and regulated by each specific state. However, with the passage of NCLB in 2002, the federal government powerfully invested themselves in public school accountability.

Intending to strengthen Title I accountability, NCLB (2002) required states to implement a statewide accountability system that assessed the performance of all public schools and students. Under NCLB all students, by 2013-2014, are expected to attain specific proficiency levels in mathematics and reading and language arts. All students can be defined as: students, regardless of race, ethnicity, socioeconomic status, gender, disability, or English proficiency (CTB/McGraw-Hill, 2002; National Association of State Boards of Education, 2002).

As a means of measuring if each student meets academic proficiency, NCLB (2002) requires both schools and districts, to make AYP. NCLB quantifies AYP by utilizing AMO. AMO can be described as the proportion of students in a specific group that meet or exceed the proficient level on a standardized assessment. Per NCLB, a school or district can be only be classified as to meeting AYP if the entire student population, as well as each specific subgroup mentioned above, are able to meet the required AMO. Academic institutions may experience potential consequences if they are unable to meet the AYP requirements for two or more consecutive years (National Association of State Boards of Education, 2002).

Unlike previous reauthorizations of ESEA, provisions within NCLB (2002) became effective in January, 2002 upon enactment of the law (Marion et al., 2002). Thus, NCLB mandated that provisions be put in place, by states, for the beginning of the current school year. It was clear that state and local educational institutions faced potential complications under NCLB; however, the urgency of NCLB left these institutions unsure of what the requirements

would be.

It was not until months after the enactment of NCLB (2002), that the USDOE gave guidance meant to clarify the expectations of states in regard to standards, assessment, and accountability. Furthermore, it was unclear how states would provide proof that they were in compliance with the requirements set forth in NCLB. According to Erpenbach and Forte (2005), states were very interested in learning more about the requirements that NCLB demanded for accountability purposes. NCLB called for a set of standards and assessments that were more rigorous than what was already in place; however, many states viewed them as an extension of the previous requirement. There were very few "new ideas" in regard to standards and assessments. More confusion arose as states developed accountability systems that met NCLB requirements all while the USDOE was still in the process of creating their regulations. With only two months remaining before states were required to submit their accountability plans for federal review, the USDOE released the final requirements (Erpenbach & Forte, 2005).

Without a clear understanding of the requirements, accountability plans varied from state to state. There was even confusion about specific strategies within states. For example, (Erpenbach & Forte, 2005) stated that the calculation of AYP indicators was not clearly explained in many of the state accountability models. Of particular concern was the varying definition of numerators and denominators in the plans. Even with the differing strategies and confusion, President Bush announced on June 10, 2003, that all of the state accountability plans submitted for review had been approved.

When NCLB (2002) took effect, many of the states were still working toward meeting the requirements set forth in IASA (1994). In fact, there were only approximately 17 states that had been considered successful in meeting the IASA standards and assessment requirements.

Many of the states were still finalizing their academic content standards and aligned assessments that were to be in place for Grades 3-5, 6-9, and 10-12. With required submission to the USDOE, states were scrambling to create detailed timelines for the development and implementation of the additional provisions under NCLB.

The USDOE did not, directly, review individual state achievement standards; however, the accountability framework introduced by NCLB (2002) was new. The new accountability framework changed the context in which achievement standards were applied. States were given the flexibility to set the bar for academic performance goals. By utilizing the percentage of students meeting this pre-determined level of academic proficiency under AYP, where a state sets the proficiency bar could impact whether or not a school or district meets the necessary requirements (Erpenbach & Forte, 2005). During the development stage, according to Hoff (2002), three states redefined their requirements of what was important for students to know and how they were able to be considered proficient. Thus, changing their definition of proficiency. As stated above, under the reauthorization of ESEA in 1994 (IASA) a completely different accountability model was being utilized by states as they were setting their academic standards.

As discussed above, NCLB (2002) focused on the amount of students reaching AYP from year to year. Under NCLB, states did not have many options when selecting indicators that were used to calculate AYP. There were five indicators that states were required to use:

- Separate summary indicators for proficiency in reading or language arts,
- Separate summary indicators for proficiency in mathematics,
- Separate indicators of participation in reading or language arts assessments,
- Separate indicators of participation in mathematics, and
- At least one other academic indicators at the elementary and middle school levels and

at least graduation rate at the high school level. (NCLB, 2002)

The graduation rate at the high school level was intended to be narrowly defined. States were able to submit an alternative to the Secretary of Education's definition (Erpenbach & Forte, 2005). At the elementary and middle school levels, states were able to choose the final indicator. States also had the freedom to include additional indicators that worked in unison with the original five indicators. According to Erpenbach and Forte (2005), few states chose to add extra indicators to their AYP model. NCLB did not define the AYP indicators as being independent or dependent of each other. With that in mind, states varied in their calculations of AYP depending on their view of the grouping of indicators. For example, some states identified school or districts for improvement if they failed to meet the AYP target for one specific indicator two years in a row. Figure 1 shows an example on how states could differ in their interpretation of AYP indicators.

		Pattern 1: The 2 indicators within a content area are independent		Pattern 2: The 2 indicators within a content area are paired			
		Reading % Proficient	Reading Participation Rate	Math % Proficient	Math Participation Rate	Other Academic Indicator	AYP Outcome
State A – only identifies for improvement using pattern 1	Year 1	X		X			In need of improvement: Reading only
	Year 2	X			X	X	
State B – identifies for improvement using patterns 1 and 2	Year 1	X		X			In need of improvement: Both Reading and Math
	Year 2	X			X	X	

Figure 1. State differences in AYP indicator interpretation. From *Statewide educational accountability under NCLB. Accountability systems and reporting state collaborative on assessment and student standards paper* by W. Erpenbach, E. Forte-Fast, & A. Potts, 2003, p. 17. Copyright 1993 by Council of Chief State School Officers.

AYP determinations were required to be made prior to the beginning of each school year. As part of the consequences under NCLB (2002), eligible parents and students were given the option of public school choice or supplemental educational services if they attended a school that failed to make AYP. For some states, the assessment used to measure proficiency was given in late spring testing. Receiving the achievement data in the summer and notifying parents prior to school year was a challenge. In these situations, some states decided to use preliminary data to identify school or district accountability placements. The USDOE encouraged states to alter their yearly assessment windows, along with make changes to their scoring and reporting windows (Erpenbach et al., 2003). These changes could have impacted the decisions states needed to make in regard to the type and quality of assessments given to students.

Criticisms of NCLB

Along with confusion about the implementation of NCLB (2002) and the effect it would have on state assessments, NCLB also gained specific criticism for its dependence on the utilization of AYP as the sole indicator school effectiveness. There are some researchers that maintain that AYP is not a true measure when examining the performance of schools (Bryk, Thum, Easton, & Luppescu, 1998; Choi, Goldschmidt, & Yamashiro, 2006; Choi, Seltzer, Herman, & Yamashiro, 2004.) AYP is an example of a status based accountability system that compares a subgroup's level or school's level of student proficiency at one point in time to an existing target level. Doran and Izumi (2004) detailed the disadvantages of using such basic measures.

Dating back to the 1960s, research of public schools in the United States has continued to find that most of the variance in student achievement among different schools could be attributed to the socioeconomic level of the school population (Teddlie, Reynolds, & Sammons, 2000).

Doran and Izumi (2004) argued that AYP based on the basic statistical measures described above provide misleading and invalid results that could also be considered a biased assessment. Schools servicing a large population of poor students are put in an unfavorable position under the AYP model. Even if students that come from high poverty families are exposed to the same rigorous education as students from highly affluent families, research suggests that they will continue to be less successful than their peers on standardized exams (Doran & Izumi, 2004).

Another disadvantage to the utilization of AYP is that it predominantly focuses on students who are right above or below the pre-determined proficiency level. A school's ability to move this subgroup of students' proficiency level does have an impact on AYP designation; however, AYP does not take into account students far below the proficiency level showing growth or students above the proficiency level pushing even higher. According to Doran and Izumi (2004), the instructional needs of the middle-level group of students can often take precedence over the needs of high-performing and low-performing students.

Finally, AYP provisions are believed to lead to an inadequate allocation of high quality educators. "Attaching consequences to test results creates obvious incentives for teachers to avoid schools that are likely to attain poor results" (J. E. Ryan, 2004, p. 974). The criticism comes from the assumption that high quality teachers would choose to work at schools AYP is more easily attained, and schools that struggle to meet AYP would "have to accept whoever is left" (J. E. Ryan, 2004, p. 974). This choice would leave students at low-performing schools, which research has connected to low socioeconomic factors, without quality educators which further creates a disparity in the achievement gap.

In response to the above criticisms, along with others, the USDOE chose to allow for some flexibility to the provisions of NCLB (2002). This study focused on the flexibility that

allows for different accountability models to be utilized, rather than the sole focus on the AYP status model.

Types of Accountability Models

Status Model

A status model is a type of accountability model that that compares a subgroup's level or school's level of student proficiency at one point in time to an existing target level, as reflected in Figure 2. As stated above, AYP under NCLB (2002) is an example of a status model that is utilized to determine school accountability. In AYP, the established target is the AMO. Accountability is measured based on the percentage of students reaching the specified level of achievement for one specific year. The status model also analyzes whether the specific subgroups of students did, or did not, meet the expected goal.

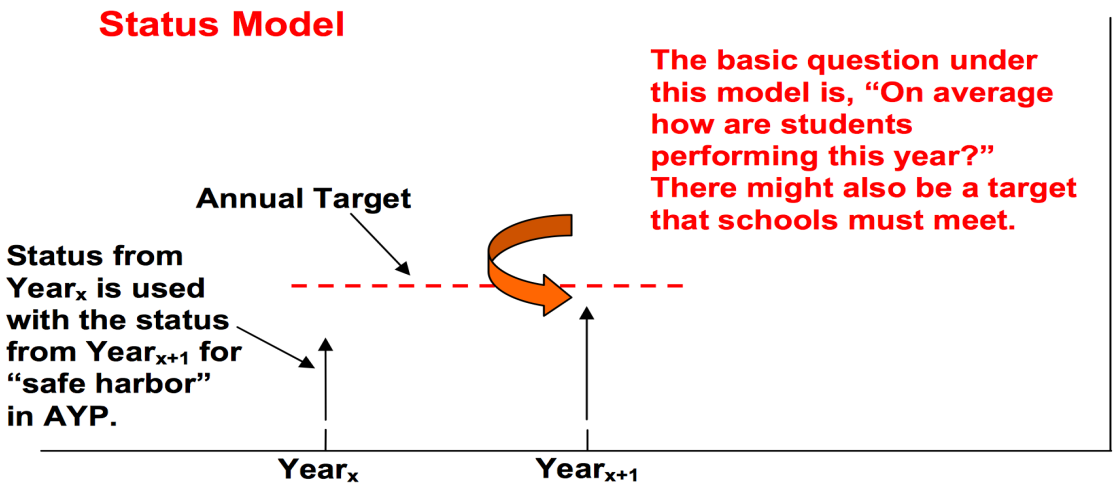

Figure 2. Status model. From *Policymakers' guide to growth models for school accountability: How do accountability models differ?* by P. Goldschmidt, P. Roschewski, K. Choi, W. Autry, R. Blank, & A. Williams, 2005, p. 3. Copyright 2005 by Council of Chief State School Officers.

The status model was widely accepted because it provided the public with an easy to understand snapshot of a subgroup or school's academic performance at a given time. However,

the criticisms of the AYP model explained above have led to further exploration of what data should be collected, as well as additional models that exist that could better assess student achievement.

Improvement Model

As explained above, NCLB utilizes an identification of subgroups within a school to monitor student achievement. The improvement model was an attempt, by NCLB, to address subgroups of students that have shown struggles in the past, such as English language learning (ELL) students or students within specific ethnic/racial groups. This type of accountability model focused on students from low-socioeconomic backgrounds, students with limited English proficiency, race/ethnicity, and students with disabilities by requiring separate reporting of assessment results from each subgroup (Linn, 2005). The reporting of these disaggregated groups was an effort to identify historically low-performing groups of students to ensure that progress was monitored throughout the school year, thus, closing the achievement gap between all students.

However, with the creation of additional subgroups, a school or district also increases additional opportunities to fail to meet AYP. The school or district as a whole could miss the AYP goal if a specific subgroup fails. Schools that service a large population of students with special needs, diverse racial/ethnic groups, or ELL students, usually considered urban schools, are placed in a potential disadvantage.

> For example, in 2003, more than 75% of the 937 California schools reporting results for two subgroups made their AYP target. In the same year, only 39% of the 291 schools reporting results for six subgroups made their AYP target. (Linn, 2005, p. 7)

In an attempt to address these concerns, the *Safe Harbor* provision was included in

NCLB (2002). Under *Safe Harbor,* a school may still meet AYP if a specific subgroup fails to meet their achievement goal. A school can meet *Safe Harbor* requirements if (a) a school decreases the percentage of students that score below proficiency by 10% from the year prior, and (b) a subgroup is able to show improvement on other indicators (Linn, 2005). It is here that the improvement model was utilized.

A status model compared student achievement scores for a specific year to an established proficiency target (Figure 3). However, the comparison of status scores at two points in time can be used to measure improvement. The improvement model is an example of how the status model was transformed to address the received criticisms mentioned above. An improvement model is a modification of the status model that was utilized to assess change between different groups of students (Goldschmidt et al., 2005). For example, the performance of a group of 7th-grade students from the current year would be compared to 7th-grade students from the year prior. Figure 5 provides a visual representation of the improvement model.

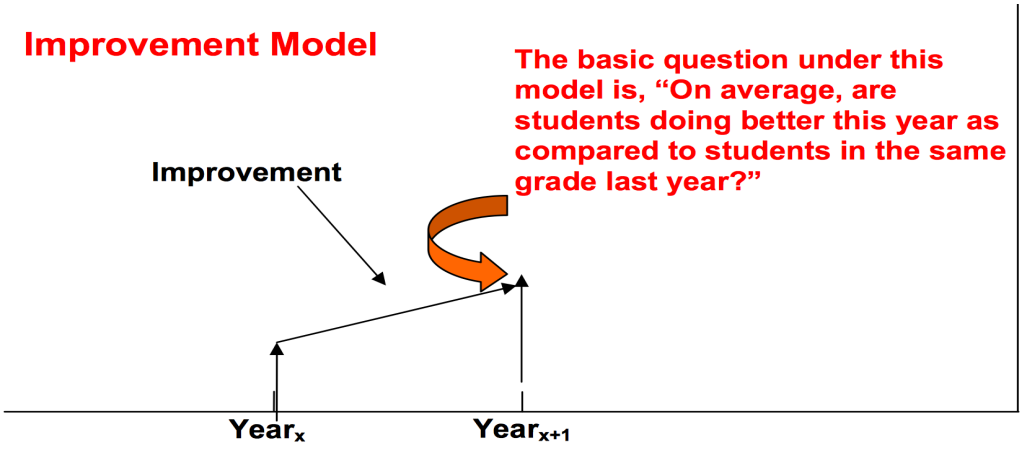

Figure 3. Improvement model. From *Policymakers' guide to growth models for school accountability: How do accountability models differ?* by P. Goldschmidt, P. Roschewski, K. Choi, W. Autry, R. Blank, & A. Williams, 2005, p. 4. Copyright 2005 by Council of Chief State School Officers.

Growth Model

A main criticism of the improvement model was that the data used to hold schools accountable consisted of different groups of students. The improvement model did not consider the potential differences among students from year to year. This included differences in socioeconomic makeup, an increase or decrease in students with disabilities, or simply, students entering a specific grade with diverse prior achievement. In other words, the improvement model was comparing two completely different groups with a specific set of data. In response, the growth model of accountability was established.

An accountability model that measures the progress of the same students from year to year is often considered a growth model. Student progress, or lack thereof, is typically observed by tracking the achievement scores of an individual student or cluster of students from one year to the next (Goldschmidt et al., 2005). For example, instead of comparing the current 7th-grade proficiency level to the previous year's 7th-grade proficiency level, the growth model compared the 7th-grade proficiency level to the same group's proficiency level in 8th grade (Figure 4).

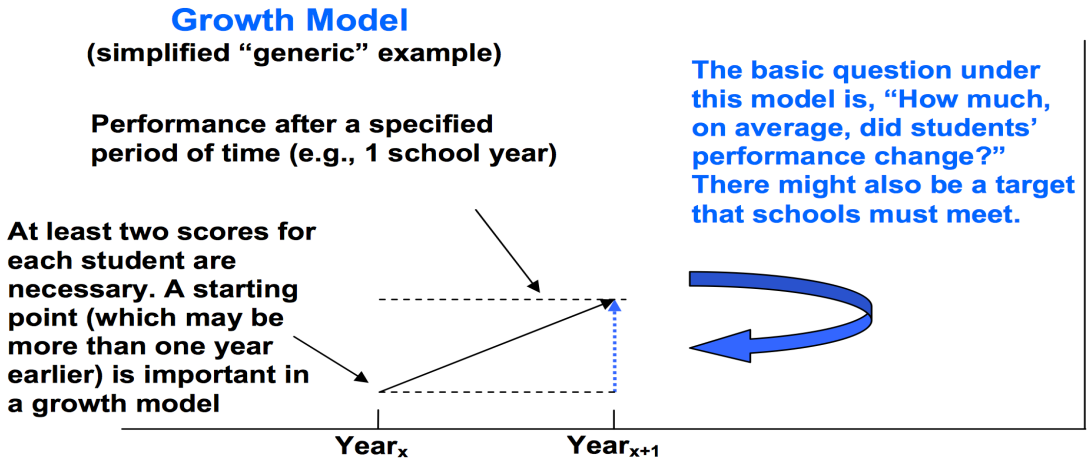

Figure 4. Growth model. From *Policymakers' guide to growth models for school accountability: How do accountability models differ?* by P. Goldschmidt, P. Roschewski, K. Choi, W. Autry, R. Blank, & A. Williams, 2005, p. 4. Copyright 2005 by Council of Chief State School Officers.

Within the growth model, individual student growth over time is compiled as a whole to identify school or district growth. According to Goldschmidt et al. (2005), progress can be defined as the degree to which a student's growth compares to the targeted growth. This can only be accomplished by observing student data over time. Rather than measuring student performance at a single point in time, a growth model assumes that student performance and in essence, school performance, over time is a more reliable indicator of academic progress (Goldschmidt et al., 2005).

Until 2005, growth models were not commonly accepted by the USDOE; however, as support continued to grow for alternative accountability models, the U.S. Secretary of Education, Margaret Spellings, revealed a new initiative that allowed for growth models to be utilized. Under this program, 10 states were granted permission to make AYP determinations, for the 2005 – 2006 school year, via growth models (Riddle, 2006). Finalized in March 2006, the initial phase of the pilot program called for allowed for eight states to "have their proposals evaluated by a national peer review panel" (Thompson Publishing Group, 2006). Later that year, Tennessee and North Carolina were granted approval, through the review panel, to implement their growth model proposals in order to make AYP determinations. In a 2010 survey conducted by Council of Chief State School Officers, 17 of the 43 responding States indicated the implementation of a growth model as their current accountability model (Klau & Autry, 2010). Furthermore, 24 of the remaining 26 responding states were identified as either considering or developing a growth model for their accountability model (Klau & Autry, 2010).

It should be noted that growth models did not replace the NCLB provisions and must meet a specific list of requirements, some of which include

- the goal of 100% proficiency by the school year 2013–2014;

- annual achievement goals should be based on the goal of 100% proficiency, not on pupil backgrounds, school characteristics, past performance or a "typical" growth rate;
- achievement gaps among subgroups should be narrowed in order for schools or LEAs to make AYP;
- progress of individual students must be tracked within a state data system; and
- assessment results must be consistent from grade to grade and year to year. (Riddle, 2006; Yu, 2007)

According to Goldschmidt et al. (2005), the design of the growth model can vary. The positive or negative relationship between growth and status should be accounted for under the growth model framework. The more years used to evaluate growth, the less initial performance will be related to growth (Goldschmidt, 2004). In other words, growth is less related to school performance based off of collective subgroup student achievement and more on the individual student performance.

Value-Added Model

Within the group of growth models, value-added models (VAM) are often commonly referenced. A VAM is an example of a growth model in which student background, prior achievement, and other data are observed to identify the specific effects of a school, teacher, or program on student academic progress (Goldschmidt et al., 2005). This type of model attempts to separate the impact of external factors, such as gender, race, socio-economic level and individual influence from a school's performance. Thus, the sole indicator of accountability is the individual student performance. The value-added score for a school is measured by calculating the difference between the actual growth and the expected growth (Figure 5).

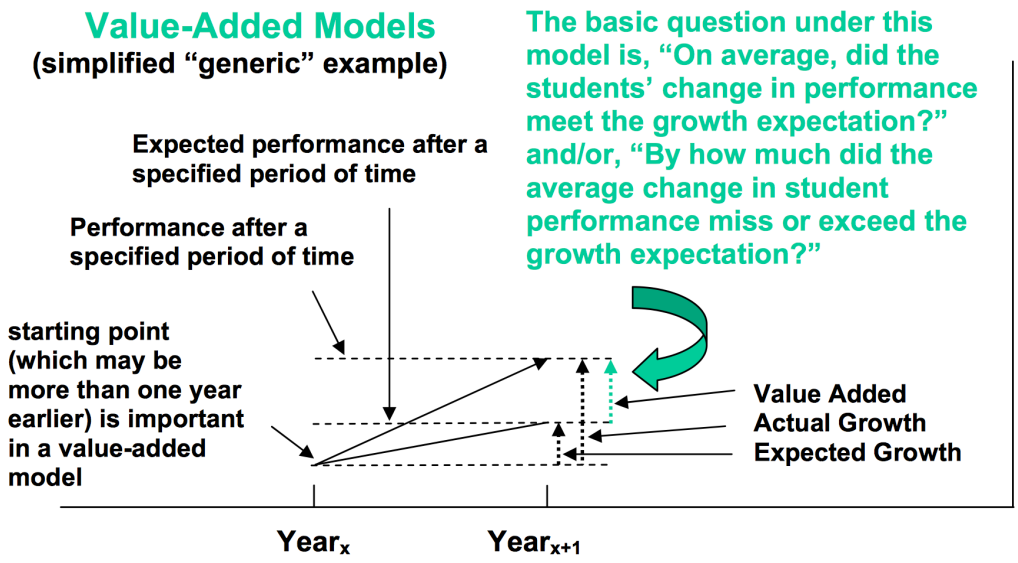

Figure 5. Value-added growth model. From Policymakers' guide to growth models for school accountability: How do accountability models differ? by P. Goldschmidt, P. Roschewski, K. Choi, W. Autry, R. Blank, & A. Williams, 2005, p. 5. Copyright 2005 by Council of Chief State School Officers.

The Tennessee Value-Added Assessment System (TVAAS) is a recognized example of a VAM. Similar to other growth models, the TVAAS system tracks the growth of student learning on a yearly basis (Sanders & Horn, 1998). What sets this apart from the traditional growth model is that the TVAAS collects data on the improvement of each student, the teacher that each student had, and the specific school that each student attended. Thus, the effect that a particular school or teacher has on student achievement is highlighted.

The ability to pinpoint the effects of a specific teacher or school has prompted many to look at a VAM as an opportunity to connect student achievement to accountability. Some districts, such as New York City, have even begun using a VAM to bridge student performance to educator pay. Advocates for VAMs argue that this could shed new light on individual teacher accountability; however, differences in opinion and research exist (Goldschmidt et al., 2005). The data received from value-added growth models could potentially lead to identifying the

effects of a classroom or a school. Less agreement exists when utilizing a value-added model to accurately measure the effect of a specific teacher (Ligon, 2008).

State Specific Accountability Models

Colorado

In February of 2011, the White House announced a list of states that were selected to receive a waiver from NCLB. Colorado was one of the a few states chosen. With this waiver, Colorado was given the authority to remove key NCLB requirements from their current accountability system. It was determined that Colorado's state accountability system already met many of the requirements set forth in NCLB. For example, the Colorado accountability model already included provisions that required schools and districts to measure performance goal progress on an annual basis.

Following the announcement that the USDOE would be allowing states to apply for a waiver to NCLB, the Colorado Department of Education (CDOE; 2011) quickly decided to pursue this option. Viewed by local officials, the waiver was an affirmation to the state's reform efforts that focused on college and career readiness (CDOE, 2011). For Colorado, utilizing student growth to measure school accountability was a critical component of the waiver flexibility, especially with regard to historically disadvantaged subgroups of students (CDOE, 2012).

Under NCLB, Colorado utilized both a state and federal accountability system. With the USDOE approving Colorado's waiver, schools and districts were being held accountable under one, unified system. This allowed for individual school and district improvement efforts to be better streamlined. More importantly, the utilization of one accountability system made it easier for the public to truly understand how schools were being held accountable.

Moving quickly, the state department of education began planning the implementation of their waiver flexibility. During this planning process, minor adjustments were made to Colorado's Educational Accountability Act in order to meet specific requirements that allowed the waiver to be accepted by the federal government.

The final in Colorado's accountability system went into effect for the 2012-2013 school year. Accountability determinations, and possible sanctions, were based on assessment results from the 2011-2012 school year (CDOE, 2011). In August 2012, school accountability was measured utilizing Colorado's self-created accountability system and the consequences associated with AYP were no longer in place (CDOE, 2012).

The Colorado Achievement Plan of Kids Act of 2008 aligned the public education system from preschool through postsecondary and workforce readiness (CDOE, 2013). The intent of this alignment was to ensure that all students graduate high school ready for postsecondary and workforce success (CDOE, 2013). The Education Accountability Act of 2009 aligned the state's education accountability framework to emphasize the objectives of the Colorado Achievement Plan for Kids Act of 2008: creating a transparent system that utilizes a set of consistent, objective measures that generates public understanding with the reporting of performance (CDOE, 2013).

The CDOE (2011) reviews each district's performance annually. In reviewing the district's performance, the CDOE compares the district's results utilizing the district performance framework that was included in the waiver application. The Education Accountability Act identified four performance indicators that would be utilized to assess a district's achievement:

- Academic Achievement: The Academic Achievement Indicator assesses the students' ability to meet the state's proficiency goal at the district level. In other words, this

indicator measured the percentage of students meeting the proficient or advanced level on Colorado's state assessment.

- Academic Growth: The Academic Growth Indicator assesses the academic progress of student proficiency utilizing the Colorado Growth Model. Growth is measured by utilizing; (1) normative growth: how growth of individual students compares to other students statewide, and (2) adequate growth: was the amount of growth comparable to that of a typical student.
- Academic Growth Gaps: The Academic Growth Gap Indicator assesses historically disadvantaged student groups and their academic proficiency and progress and struggling learners. The Growth Indicator is separated into student groups and compares each group's growth with expected normative growth.
- Postsecondary and Workforce Readiness: The Postsecondary and Workforce Readiness Indicator analyzes the school or district's ability to adequately prepare students for college or career opportunities post high school. (CDOE, 2014, p. 8)

State-identified measures and calculations are used to give a rating on each performance indicator to the district. Districts are given a specific determination if they exceeded, met, approached or did not meet the state's expectations. Each indicator determination is then combined to create an overall district assessment. Districts can be placed within five different categories. These categories are

- Accredited with Distinction – state attainment expectations for the Performance Indicators are met or exceeded by the district. Districts that meet this distinction are required to adopt and implement a Performance Plan;
- Accredited – state attainment expectations for the Performance Indicators are met by

the district. Districts that meet this distinction are required to adopt and implement a Performance Plan;

- Accredited with Improvement Plan – the district has not met state attainment expectations on the Performance Indicators. Districts are required to adopt and implement an Improvement plan;
- Accredited with Priority Improvement Plan – the district has not met state attainment expectations on the Performance Indicators. Districts are required to adopt and implement a Priority Improvement plan; and
- Accredited with Turnaround Plan – the district has not met state attainment expectations for the Performance Indicators. With approval from the commissioner, districts are required to adopt and implement a Turnaround plan. (CDOE, 2014, p. 9)

The NCLB Flexibility waiver replaced the previous district Title I Accountability measure, AYP, with Colorado's district performance frameworks. Districts now receive one set of accountability data for both Title I and state accountability.

The CDOE is responsible for developing and maintaining a Web portal to provide information about student, school, and state performance to public schools, school districts, the Charter School Institute, parents, and other members of the public. The Web portal provides performance reports for schools, districts, and the state, as well as accreditation categories and school plans.

Tennessee

In Tennessee, a large shift in educational improvement began in the 1980s. With the hopes of improving the educational opportunities for students, Governor Lamar Alexander enacted the Comprehensive Education Reform Act (CERA) in 1984. Included in CERA was a

significant increase in educational spending and the creation of a career pathway for teachers (Sanders & Horn, 1998). This was Tennessee's first step in implementing a teacher evaluation system that included a pay for performance.

In 1989, Governor Ned McWherter signed the second major piece of academic legislation into law by signing the Education Improvement Act. Legislators representing both parties adamantly sought out a provision to the framework that would require schools and districts to utilize the new financial support to improve student academic achievement (Sanders & Horn, 1998). As mentioned earlier in this chapter, this new accountability provision would become known as the TVASS. To create the state's new accountability system, data in regard to attendance, promotion, graduation and dropout rates of individual schools was collected (Sanders & Horn, 1998).

The TVASS created accountability determinations for districts, schools, and teachers based on the growth of individual student academic achievement over time. According to Sanders and Horn (1998),

> an integral part of TVAAS is a massive, longitudinally merged database linking students and student outcomes to the schools and systems in which they are enrolled and to the teachers to whom they are assigned as they transition from grade to grade. (p. 1)

Contrary to other opinions, Sanders, a statistician from the University of Tennessee, stated that "research conducted utilizing data from the TVAAS database showed that race, socioeconomic status, class size and classroom heterogeneity are poor predictors of student academic growth" (Sanders & Horn, 1998, p. 1). Rather, Sanders and Horn (1998) argued that the "effectiveness of the teacher is the major impact on student academic progress" (p. 1).

The TVAAS is a statistical based accountability model that calculates the effectiveness of

districts, schools, and teachers. "TVAAS uses statistical mixed model theory and methodology to enable a multivariate, longitudinal analysis of student achievement data" (Sanders & Horn, 1998, p. 1). Included are the individual student proficiency scores on (a) the Tennessee Comprehensive Assessment Program and (b) end-of-course tests in high school subjects. Collected over time, individual student achievement data is linked to that student's teacher(s), school(s), and school system(s). By looking at the data over an extended period of time, TVAAS is able to generate an individual student-learning pattern for each student in Tennessee. This student-learning pattern allows individuals to observe academic growth or deviation for individual students.

The design of TVAAS was centered around the framework's ability to provide data needed to evaluate the effectiveness of a school, along with the teacher, in regard to student growth in academic proficiency over a three-year period (Sanders & Horn, 1998). TVAAS provides districts, schools, and teachers detailed reports outlining the impact each had on the achievement of students. These reports allowed educators the information needed to effectively plan, alter methodology, and provide programs needed to effect student growth. According to Sanders and Horn (1998), this information "has been found to be invaluable by many teachers and school administrators involved in the curricular planning, program evaluation, and developing strategies to meet the needs of student with differing academic attributes and abilities" (p. 3).

Two critical components of TVAAS were the use of multiple assessment measures, including the Tennessee state-level assessment and individual district-level assessments, and shifted the focus from single year proficiency to using multiple years to determine peer group comparative growth percentiles (Fujino, 2013). Contrary to earlier trends in educational

accountability, TVAAS shifted the conversations from the complexity of state assessments, to a discussion about changing the interpretation of the student achievement data gathered through this assessment tool (Sanders & Horn, 1998).

Indiana

The standards and accountability reform movement in Indiana began in 1987. Then Governor Robert Orr, along with State Superintendent of Public Instruction, H. Dean Evans pressed the Indiana General Assembly to create the A+ Program. This initial legislation produced an accountability system that generated awards and accreditation based on student performance. The implementation of the Indiana Statewide Testing for Educational Progress (ISTEP) program provided the student performance data required to determine school accountability (Gold, 1988).

During the late 1990s, the Indiana General Assembly was presented with various educational reform efforts. Most of these efforts failed to pass prior to the end of the legislative session. In 1999, new legislature, supported by the Indiana Chamber of Commerce, was presented to the General Assembly. The accountability reform system that was presented became Indiana Public Law 146 (HEA 1001-2008) and Indiana Public Law 221 (HEA 1750-1999). P.L. 146 called for the development of rigorous academic standards by IDOE and for the adoption of these standards by the State Board of Education (SBOE). These standards were to be developed English, mathematics, science, and social studies from Kindergarten through Grade 12. Furthermore, P.L. 146 required that standards be developed for subjects not tested by ISTEP. P.L. 221 was passed in conjunction with P.L. 146 to create an annually reported accreditation and accountability system that was focused on student performance. P.L. 146 also provided financial incentives for schools determined to be high-performing, as well as funding for

professional development. According to Hiller, DiTommaso, and Plucker (2012), a core component of P.L. 221 was the requirement to adopt an annually revised, three-year school improvement plan.

Working in close collaboration, the IDOE and the Education Roundtable took two years to design and write the provisions of P.L. 221. Throughout this two-year process, disagreements did arise concerning the actual implementation of the system passed by the legislature (Hiller et al., 2012). Of major concern was the school improvement labeling system that P.L. 221 utilized to categorize schools. After much discussion, a compromise was reached to create the accountability labels for schools. Under P.L. 221, schools were designated into the following categories: exemplary progress, commendable progress, academic progress, academic watch, and academic probation (Zehr, 2001).

The design of P.L. 221 took place prior to NCLB (2002); however, NCLB mandated that the federal accountability system be embedded within each state's own accountability system. For Indiana, this was a challenge. Specific differences between NCLB's AYP system and Indiana's P.L. 221 existed in regard to the measurement of student learning. First, AYP only measured the percentage of students passing ISTEP, and P.L. 221 incorporated improvement, as well as performance. Second, AYP only utilizes comparisons of the same grade level on a year-to-year basis; whereas, P.L. 221 focused on student cohort groups from year-to-year. Finally, P.L. 221 measured student performance for all students tested, and AYP only utilized student performance data for student demographic subgroups (Hiller et al., 2012).

To make this accommodation, P.L. 221 was written to prohibit any school to place higher than the academic progress category if, for two consecutive years, they failed to meet AYP. This created further difficulty for schools that were truly trying to improve as they were actually

working to adhere to both, and often competing, pieces of legislature.

Following the NCLB waiver announcement by President Obama, Indiana quickly applied due to the solid accountability framework already in place. During a December, 2011, SBOE meeting, former State Superintendent of Public Instruction, Tony Bennett, expressed his support of the waiver application, stating that Indiana's accountability system was "about as solid as anyone [*sic*] in the country" (Bennett, 2011, 9:07am). Indiana's waiver application, along with Colorado, was one of 10 states chosen to receive the AYP exemption waiver. This exemption allowed Indiana to no longer utilize AYP in calculating category placement (Elliot, 2012).

Turning to a new accountability model, Indiana utilized its own growth model instead of a status model. The Indiana A–F Accountability Model was put into place in 2011. Each school in the state of Indiana would receive an A–F accountability grade. Bennett (2011) argued parents, students, and educators in Indiana needed a more understandable reporting system showing the progress and status of their schools. For the elementary and middle levels, grades are determined by three categories: performance and improvement, growth, and participation.

- Performance and Improvement – Schools receive a score based on the percentage of students passing the math and English/language arts portions of ISTEP.
- Growth – The Indiana A–F accountability model is considered a growth accountability model. This system assesses the academic proficiency progress each student makes in a specific year. The growth on ISTEP, high or low, that students demonstrate will be used to increase or decrease a school's overall score.
- Participation – Schools are required to meet a specific percentage of student participation on ISTEP. If schools do not meet this threshold, their score is lowered. (IDOE, n.d., p. 1)

Variables That Impact Student Achievement

Content and Performance Standards

Much of the national reform efforts have been centered on content and performance standards. The creation of content standards was an attempt to define the knowledge and skills that all students are expected to know (Duran, 2005). The performance standards define the level of student knowledge. This is often described as advanced, proficient, or basic. The proficient level, at minimum, is the expectation for all students. Schools and districts are measured based on the percentage of students that improve, or move, into different ability categories.

Duran (2005) contended that "requiring the same content and performance standards creates a number of concerns because all students are expected to achieve, at minimum, a proficient level of academic ability at benchmark grades" (p. 81). This achievement by students has the potential to be very difficult as not all students are able to learn the same curriculum in the same amount of time. According to Goertz and Duffy (2001), there are unequal opportunities for children, in the United States, to learn and succeed. Linn (2000) argued that students begin school, at a very early age, with different level of academic abilities. Although still having high standards for all students, Linn suggested not having common standards for all students. When considering students with special needs and ELL students, high performance standards does not necessarily have to mean having the same academic standards for all students at each grade level.

State Standards Assessments

As states create their student assessments as a means to measure student achievement, norm-referenced tests are often the type selected. Norm-reference tests provide quick and easy

data that allow states the ability to compare results, make decisions about instruction, and to assess learning. Norm-referenced assessments are designed to ensure that half of the student scores would fall at or below the normal distribution of scores and that the remaining half of student scores would fall at or above the normal distribution. This type of assessment creates a bell-curve plotting of results.

Standards-based assessments also have the ability to serve as indicators of student and school performance. Some states have attached incentives and consequences to student achievement data with the intent of encouraging teacher and student performance (Baker, 2002). This utilization of assessment data to provide consequences and/or rewards is regularly used today, despite the primary purpose of assessments being the opportunity to gather crucial data that can be used to impact student learning and instruction. As assessments continue to be used for purposes outside their original intent, new concerns arise about the usefulness of the information (Duran, 2005).

Interpretation of Results

Interpreting data, both clearly and accurately, is a difficult task (Duran, 2005). Standardized test results could be utilized, by the public, to create perceptions about an educational institution. This ranking of schools and districts is not always fair, as there are many other factors that determine an effective school (Linn & Haug, 2002). A school or district's quality is interpreted by their changes in assessment results (Duran, 2005). All too common, these changes, either positive or negative, in achievement scores are attributed to the quality of instruction taking place (Baker, 2002). However, these differences can occur due to various reasons, which include students moving in/out, an increase/decrease in specific student groups, redistricting, etc.

Another common practice for states is to compare the changes of yearly assessment results for a school or district. Statistically speaking, there are potential variables that are not controlled for when interpreting results in this manner and can lead to further confusion about the actual results (Linn, 1999). This type of comparison could be considered acceptable for schools or districts that have little turnover in student population. However, the validity of inferences for schools with transient demographic groups or with schools that have a small population of students in specific grade is questionable (Duran, 2005). Alternative interpretation methods are available that could result in more accurate accountability determinations, and academic proficiency improvement. According to Brennan (2001), an example of such an alternative method would be analyzing individual student proficiency results over a number of years to monitor student progress.

Factors outside the school's control could also lead to changes in student results (Duran, 2005). Examples of these factors include (a) the specific students tested within school, (b) statistical errors, (c) the student population size, and (d) a lack of contextual information. These items are not always readily available when interpreting the quality of a school through statistical information (K. Ryan, 2002). A family's educational background and cultural expectations are examples of contextual information that may result in a better interpretation of differentiated results (Glidden, 1999). These factors can have a monumental impact on student and school achievement.

There are a number of factors that cause student achievement to differ within and between schools. These factors reflect the challenges that each specific school or district face. Schools that lack diversity serve students of specific homogeneous groups (Duran, 2005). These homogeneous groups can contain students from specific racial and socioeconomic makeup.

However, an often-overlooked variable could be the location or setting of the school. A school or district's location could also have an impact on student achievement. Schools from rural locations could possibly have a very different school culture than schools located within the inner-city. Watts (2012) grew up in a small Kentucky town as a member of a Black minority. He later returned to his hometown to become a professor in his town's community college. Watts firmly believed that small towns needed to support the education of minority students. Through his experiences, Watts pushed for schools to develop an inclusive educational setting for students. Watts found that student achievement for minority students improved in this inclusive setting, even if minority populations were isolated in the community itself (Riggen, 2013; Watts, 2012). Many small towns and rural communities continue to face challenges as immigration and the push for cultural competency grows (Villenas, 2001).

School Size

Of particular interest to educators in the United States, is the relationship of school size and student achievement (USDOE, 2009). The origin of education began utilizing small, one-room school houses that were often set in rural areas. However, in the early 20th century, many reformers stated that these types of school settings were "ineffective, inefficient, and hindered by provincial attitudes and local politics" (Arnold, 2000, p. 3). According to Fowler and Walberg (1991), consolidation quickly became the preferred solution. Between 1930 and 2000, the U.S. student population increased 83%, and the number of U.S. school districts decreased 91% (Crouch, 2012). This was an indication that larger districts were being created. Specifically referencing Indiana's School Corporation Reorganization Act of 1959, the number of districts in the state decreased from 900 to 400 in a 10-year time period until 2010, when there were approximately 292 school districts (Dokoupil, 2010; Zimmer, 2007). School reorganization

followed the industrialization trend, in which an increase to the size of the organization, or facility, could lead to a significant financial reduction. Increasing school size also addressed educational quality concerns for small, rural schools. In his 1959 book *The American High School Today*, author James Bryant Conant argued that "many American public schools were too small to allow diversified curriculum and equal educational opportunities for their students" (p. 77).

Many research studies exist that attempt to draw relationships among school size and student achievement. Research conducted by Bancroft, Barker, and Gump (1964) found that student involvement and high school size had a significant relationship. Some of the benefits of attending a smaller school included an increased student participation rate in school activities, more opportunities for students to hold leadership positions, and an increase in school/community identification. Bancroft et al. (1964) argued that these benefits all lead to higher student achievement. When it comes to finding the optimal number, research differs. Sizer (1992), along with Gregory and Smith (1987) determined that a school with 250 or fewer students was optimal for having the greatest impact on student achievement. However, Goodlad (1984) argued that a population of 500 to 600 students was optimal.

In 2005, the Kentucky legislature requested a study to be conducted on the impact of school size on student performance. Reviews of research literature on the effect of school size, followed by an analysis of Kentucky's 1,200 public schools were conducted. Through an examination of current research, Hager (2006) identified a belief that attending larger schools negatively impacted the academic proficiency of disadvantaged students. However, Hager's research found the exact opposite. Student achievement was greater for high school students attending some of Kentucky's largest schools (Hager, 2006).

Lowen, Haley, and Burnett (2010) analyzed the impact consolidation had on creating larger schools. The benefits of consolidation were questioned, as to whether they outweighed the challenges associated with creating larger schools. Lowen et al. (2010) were concerned with the undermining of the principals' abilities to monitor teachers' efforts in the classroom. These efforts in the classroom, obviously, have an effect on student performance and achievement. It was concluded that an incentive-based system should be included to encourage teacher performance.

Being a small school, however, does not automatically result in increased student achievement (Arnold, 2000). In a report prepared published for the Institute of Education Sciences, Arnold (2000) concluded that school size was one of many variables that impacted student achievement and that smaller schools should identify ways to utilize their size to positively impact student achievement. In England and Wales, Hopkins and Ellis (1991) found that school size does, indeed, impact student learning; however, factors such as teaching methods, availability to resources, curriculum and leadership have a much larger impact.

Socioeconomic Status

The concept of consolidation has led to schools becoming much larger since the mid-1950's (Irmsher, 1997). During this time, urban schools were largely attended by students from middle-class families (Howley, 1994). However, as time has passed, there has been a major shift in the student populations of urban schools. Today, many of the large urban schools are filled with impoverished, disadvantaged students. In research conducted by Fowler and Walberg (1991), a district with low socioeconomic status was the most consistent predictor of poor student performance. A 2005 study found that the combination of ethnicity, SES, and property value led to a 70% difference in school achievement outcomes when analyzing Illinois Standards

Achievement Test scores for Grades 3-8 (Beck & Shoffstall, 2005).

Compared to their peers, children who come from families of poverty are already behind when they begin school. Jensen (2009) argued that the early childhood years are a crucial time for brain stimulation and development. Poor children, when compared to their well-off peers, received less of this stimulation (Jensen, 2009). Studies have shown that this lack of stimulation can lead to cognitive, social, and emotional difficulty, as students increase in age. Brain stimulation, or lack thereof, has been shown to effect vocabulary growth, IQ, and social skills (Bradley, Corwyn, McAdoo, & Coll, 2001). Poverty has been proven to have a direct relationship with cognitive achievement on standardized assessments. Assessment scores in reading, math, and science are often below average for students from low socioeconomic families (Jensen, 2009).

Socioeconomic inequalities among schools and districts also effect student achievement. In many instances, schools with high levels of both poverty and minority populations receive less financial support from state and local government than more advantaged schools. Research also suggests that high-poverty, high-minority schools are more likely to employ inexperienced teachers or teachers who are teaching outside their specialties (Jerald, 2001) as reflected in Table 1.

Table 1

Percentage of Educators in High Poverty Schools Teaching Outside Their Subject Expertise

Demographic	Math	English	History	Physical Science
All Public Schools	35.8%	33.1%	58.5%	59.1%
High Poverty Schools	51.4%	41.7%	61.2%	61.2%

Source. Adapted from *Dispelling the Myth Revisited: Preliminary Findings from a Nationwide Analysis of "High Flying" Schools*, by C. D. Jerald, 2001.

A study conducted by Sum and Fogg (1991) found that mid-upper income students rank in the 66th percentile on state assessments, and students from poverty are ranked in the 19th percentile. Regardless of race or ethnicity, students from low-income families consistently score below average (Bergeson, 2006). In one study, 13% of low-income students met all subject area thresholds, and 44% of the same students did not meet any of the subject area thresholds (Bergeson, 2006). Additional research found similar results. "Children who lived in persistently poor families scored 6 to 9 points lower on the various assessments than children who were never poor," (Smith, Brooks-Gunn, & Klebanov, 1997, p. 164).

Race

Research studies have shown that the race of a student is closely associated with poverty as a predictor of student achievement. Harkreader and Wathersby (1998) found that poverty had a larger impact on student achievement; however, Bankston and Caldas (1998) found that minority status impacted student achievement more than a student's socio-economic status. Data suggests that African American children (33.1%) have a higher likelihood to live in poverty than White children (13.5%). African American students are also more likely to have single parents, and more likely to be welfare dependent (Rector, Johnson, & Fagan, 2001). Furthermore, a correlation between minority status and lower teacher certification and experience exists (Darling-Hammond, 1999.)

Gender

Research studies show that, on average, female students do better in school than their male counterparts. Girls get higher grades and complete high school at a higher rate compared to boys (Jacobs, 2002). Results from standardized tests also show that female students perform better in literacy, writing, and general knowledge (U.S. Department of Education, 2001).

Disaggregation of the 1998 NAEP reading results by gender rather than race (Donahue, Voelkl, Campbell, & Mazzeo, 1999) found that girls outperformed boys in all grade levels tested. Female students continue to show higher verbal scores throughout high school, but they being to lose ground to male students after fourth grade on mathematical and science standardized tests. Teacher characteristics and the classroom environment have also been identified as variables contributing to the gender gap (Zembar & Blume, 2009).

The Principal's Impact on Schools

Research studies that attempt to identify the variables that make schools successful are nothing new. Many studies exist that examine teaching practices, school size, curriculum needs, student capacity, etc. One consistent finding is that the principal is the key to implementing and achieving comprehensive school reform (Carter, 1999; Goldberg, 2000). Commonalities in these studies found that principals had autonomy and control. Principals in turn gave their teachers much autonomy and supported them through staff development. These principals realized the importance of building relationships with staff, parents, and community but ultimately held themselves accountable for all aspects of the school (Teskie & Schneider, 1999).

Keller (1998) found that most of the principal effectiveness research had focused on the principal's role in school change and school success. His studies identified the importance of the principal's leadership role in creating and maintaining effective schools. In particular, Keller (1998) noted the importance of the principal serving as leader of curriculum and instruction.

Emphasized in recent literature, the importance of the principal and the principal's leadership style has been explored (Barth, 1990; Edmonds, 1979; Sergiovanni, 1995). Dimmock and Lee (2000) distinguished between leadership and management:

Leadership refers to higher-level practices, such as setting visions and goals, motivating

> staff to achieve them, and building culture and climate. Management, on the other hand, refers to lower-level practices such as monitoring and maintaining performance, allocating resources to achieve agreed-upon goals, and organizing the curriculum and the class schedule. (p. 339)

Setting the instructional climate and emphasis is a critical role of the instructional leader. With increased accountability being placed upon the individual schools for improving student achievement on standardized test scores, the principal's role becomes even more important. Instructional leaders set high expectations for their staff. In turn, research conducted by Rosenthal (1994) found that students lived up to teacher expectations. Students scored higher when teachers placed high academic expectations upon them.

A study of effective and ineffective elementary and high school principals by Heck, Larsen, and Marcoulides (1990) found that "through frequency and effectiveness of implementing instructional leadership behaviors identified, principals can have direct effects on the achievement levels of their schools" (p. 120). Some of these behaviors can be directly linked to the classroom, like classroom observations and providing feedback on instruction, but many of them informally impact the classroom environment like school climate and goal setting (Turner, 2013).

Research suggests that, outside teachers, school leaders are the second most significant influence of student achievement (Seashore-Louis, Leithwood, Wahlstrom, & Anderson, 2010). In just one year, highly effective principals have the potential to increase student proficiency results by 10% on standardized tests (Waters, Marzano, & McNulty, 2003). The educational culture principals create has the potential to impact attendance rates, disciplinary actions, and graduation rates, as well as other student outcomes. Research also suggests that a principal tends

to have a more significant impact on student outcomes at schools that are low-performing or have a high minority/high poverty population (Leithwood, Louis, Anderson, & Wahlstrom, 2004; Seashore-Louis et al., 2010).

Clark, Martorell, and Rockoff (2009) found that principals become more effective in their first three years. Even more important, their study found that it takes approximately five years for a principal to sufficiently impact a new school regardless of their effectiveness a their prior school. This timeframe allows principals to improve staff and implement policies that will have a positive impact on student performance (Seashore-Louis et al., 2010). This is not to say that effective principals cannot make significant improvements during their initial tenure; however, their overall effectiveness improves as their experience increases. With that in mind, research suggests that principals at the most challenging schools are more likely to have less experience leading a building (Loeb, Kalogrides, & Horng, 2010).

Summary

This chapter introduced a multitude of research in regard to school accountability in the United States. It is evident that school accountability is not a new concept. Moreover, school accountability is no longer a topic that is discussed only amongst the education profession. School accountability has taken a social shift to be involved in politics and in general population discussions.

The first section of this chapter looked at the evolution of school accountability in the United States. School accountability began as a way to assess the progress of individual schools. Information was gathered to allow schools and districts to improve curriculum, teaching practices and student opportunities to be challenged. Due to social issues within the United States, school accountability quickly morphed into a politically generated response dealing with

these issues. Education and school accountability was, originally, handled at the state level. However, over time this also shifted into a national agenda.

The second section of this chapter researched the introduction of NCLB (2002). For the first time in history, the education system became infused by the federal government with money and regulations. This significant piece of legislation was a critical force in how the government and the public looked at school accountability.

The third section of this chapter investigated multiple accountability models. The status, improvement, growth, and VAMs were introduced. With the federal focus on school accountability, research has also increased on the best way to measure for accountability. This section also looked at the evolution of accountability in three specific states. These states began using the AYP measure of school accountability framework set our under NCLB. However, as concerns continued to arise, these three states attempted to create their own accountability models. This section also looked at the journey of these states and the different directions these states chose to take when measuring school accountability.

The final section of this chapter looked at variables that affect school accountability. This section is a critical component of this paper's investigation. As states and the federal government continue to fine tune their school accountability process, it is important that they include the right variables. Even more important is that the policymakers understand the impact that certain variables have on student achievement.

CHAPTER 3

METHODOLOGY

This chapter presents the methodology of this study, beginning with a description of the purpose for which this study was conducted. This is followed by a description of the research design and the research questions and hypotheses. A description of the methods employed to analyze the collected data precedes a summary of the chapter.

As stated in Chapter 1, this study examined a portion of the NCLB accountability system, along with alternative accountability models, in an attempt to identify the variables members of the education profession believe should be utilized. This study also examined the reliability and validity of AYP, as well as other accountability variables that are often utilized to measure school quality. Recent research suggests that reliability issues exist due to the lack of consistency with utilizing AYP to measure school accountability (Elmore, 2002; Guilfoyle, 2006; Kane, Douglas, & Geppert, 2002; Linn & Haug, 2002; Porter et al., 2005; Schoen & Fusarelli, 2008; Wiley et al., 2005). Thus, this study also evaluates the variables that are utilized in other accountability models. Since the purpose of these accountability models is to improve schools and the education that students receive, the variables used to determine school accountability should assess the proper data to allow for increased student achievement and school improvement. If school improvement and student achievement are found, it could be inferred that the goals set forth in NCLB (2002) are being met. On the other hand, a conclusion

can be made that NCLB is not successful if schools fail to meet progress requirements and are not working toward improvement. Furthermore, if schools receive different accountability labels simply due to potentially irrelevant variables being used, can conclusions truly be made about whether a school is successful or not?

Improving the educational opportunities for underperforming students and schools should be the intent of accountability reform. NCLB (2002) mandates that student achievement data must be used as a means to measure accountability. This study investigated potential composite score differences between the perception of the variables used to measure school accountability and the type of position held, location of the school/district, socioeconomic make-up and the current school/district Indiana A–F accountability grade. An analysis of this data allowed for relationships to be made in regard to demographic data and the variables used to measure accountability. Instead of measuring school effectiveness, researchers (Elmore, 2002; Harris, 2007; Mathis, 2004, 2006) argue that AYP measures the demographic and socioeconomic status of the schools. The validity of AYP could be inconsistent if this is truly the case. States have spent a great deal of time exploring alternative accountability measurements. This study quantifies the contention that measuring accountability with different variables could lead to different outcomes. Data were collected from a survey that was sent to teachers, principals, and superintendents in the state of Indiana.

Research Questions

The proposed research questions were introduced in Chapter 1.

1. What are the current perceptions of Indiana educators on the variables in various accountability systems?
2. Is there a significant difference in the accountability composite score based on the

type of position of the respondents?

3. Is there a significant difference in the accountability composite score based on the school or district's accountability grade?
4. Is there a significant difference in the accountability composite score based on the school or district's socioeconomic level?
5. Is there a significant difference in the accountability composite score based on the school or district's location type?
6. Does school socioeconomic status level, school enrollment size, years in current position as a principal, location, and school accountability grade serve as predictors of the composite score?

Quantitative Research Design

This quantitative, survey-based research sought to determine teachers,' principals,' and superintendents' perceptions of the variables needed to effectively measure school accountability. Also investigated were the responses from these three groups in relation to four different criteria: (a) perceptions based on position type, (b) perceptions based on the school/district's Indiana accountability grade, (c) perceptions based on the socioeconomic make-up of the school/district, and (d) perceptions based on the school/district's location type. The quantitative design of this study allowed for statistical models to be run answering each research question.

Survey-based research was chosen for two reasons. First, the review of literature indicated that numerous qualitative-based research projects have been conducted on the concept of school accountability, but little evidence exists of survey-based research that explores the perceptions of different job roles and their beliefs on what they feel should be utilized to measure

school accountability. It is hoped that the instrument developed in this research will be useful for future research. Second, as Creswell (2009) indicated, survey-based research is an economical and advantageous way of identifying the perceptions of a large population. Additionally, survey-based research affords researchers a rapid turnaround in data collection (Creswell, 2009).

Data were collected via a Web-based survey. My sponsoring university made available to me a Web-based program for designing online surveys. Web-based surveys are more economical and efficient than paper-based surveys (Cox & Cox, 2008).

Sample

The sample addressing each of the research questions in this study included more than 63,000 teachers, 800 principals, and 280 superintendents in the state of Indiana. I utilized distribution lists obtained from the IDOE for each of these groups to disseminate the survey to each population.

Provided in this sample were schools and districts that differ in traditional AYP categories such as school size, socioeconomic makeup, racial/ethnic groups. This sample also included respondents from various school/district location types (rural, urban, and suburban). Also included in this sample size were schools that received various Indiana A–F accountability grades. Finally, this sample consisted of individuals who had differing levels of experiences within their current position

The identity of individual participants was protected via the use of the Web-based survey provider, Qualtrics; I had no way of identifying the responses of any individual. Each participant was informed of the anonymity of his or her responses as part of the invitation to participate. Procedures for administering the Web-based survey received Indiana State University

Institutional Review Board (IRB) approval before the survey was distributed. There were no risks associated with participation in this survey, as the design of this measurement tool maintained confidentiality. Respondents received no compensation for their participation in this survey. Finally, if any individual who participated in this survey wished to learn the results or would want access to the research findings, they could e-mail me to receive a copy of the findings.

Survey Design

Participants were asked to participate (Appendix A), and if they agreed, complete a survey consisting of 17 questions (Appendix B). The survey consisted of four sections. Section 1 consisted of five questions that asked participants to share their opinions of variables currently used by the state of Indiana to measure school accountability. Participants were asked to respond to the questions using a Likert scale of 1 – 6 with 6 = *strongly agree* and 1 = *strongly disagree*. The five responses were added together to serve as the composite score to answer Research Questions 2 through 5. This composite score was utilized because it represented the current perception of Indiana teachers, principals, and superintendents. Section 2 consisted of six questions that asked the participants to share their opinions about examples of variables used, outside the state of Indiana, to measure school accountability. Participants were asked to respond to the questions using a Likert scale of 1 – 6 with 6 = *strongly agree* and 1 = *strongly disagree*. Section 3 consisted of four demographic questions and two open-ended questions. Research Question 1 was answered with the utilization of descriptive statistical data, and Research Questions 2 through five utilized inferential statistical data. Section 4 of this survey consisted of three questions specific to principals only. The statistical data were used to address Research Question 6. The three questions were used to run a multiple regression analysis to look

at the strength of the relationships between the composite score variables and the predictor variables. By looking at the strength of these relationships, this study described the amount of variance amongst the criterion variables (composite scores) and the predictor variables.

Content Validity

Content validity provides confirmation that the study is measuring the concept the researcher that was attempting to assess. Creswell (2009) stated, "establishing the validity of the scores in a survey helps to identify whether an instrument might be a good one to use in survey research" (p. 149). A thorough investigation into research studies in Chapter 2 helped determine content readability. The establishment of content validity occurred through the utilization of educators from the Indiana State University doctoral cohort. Due to their positions, these participants were not a part of the survey. This group analyzed the school accountability survey questions and reviewed the questions from the school accountability survey that was created. The group presented critiques to assure face and content validity were present.

Survey Reliability

Howell et al. (2012) stated that reliability is the extent to which a study produces similar results. Without the ability to replicate the measurement procedures, researchers would find difficulty in drawing conclusions or making claims about their research. Survey reliability is the degree of stability shown when a measurement is repeated under consistent conditions. The school accountability survey results used in this study were analyzed for reliability using a Cronbach's alpha test, the most commonly used measure of reliability (Tavakol & Dennick, 2011).

Data Collection and Procedures

This research used a Web-based survey as the instrument to determine teachers,' principals,' and superintendents' perceptions about the variables used to effectively measure school accountability. The instrument was designed by me specifically for this research and was administered through the online survey service, Qualtrics. An e-mail to all Indiana public school teachers, principals, and superintendents was sent out requesting participation in this study. The IDOE was utilized to attain the e-mail addresses for all participants. An informed consent letter (Appendix A) preceded the survey (Appendix B). The study's purpose and directions for accessing the Qualtrics online survey were included in the informed consent letter. The informed consent letter also explained the confidentiality procedures that were in place to maintain anonymity. Participants were also informed that continuation of the survey could be stopped whenever they wished. Following the informed consent letter, participants could choose to take part in as much of the survey as they liked. The survey was distributed in the fall of 2014.

The questions asked teachers, principals, and superintendents their perceptions of school accountability variables. The data were analyzed to identify educators' beliefs connected to the appropriate variables used to measure school accountability. Participants had two weeks to participate in the survey. During this window, a participation request e-mail was sent to all Indiana teachers, principals, and superintendents. Upon closure of the participation window, access to the survey was unavailable and all survey responses were statistically analyzed.

Statistical Analysis

Research Question 1 was tested using descriptive analysis. Means and standard deviations were determined from the school accountability perceptions of Indiana teachers, principals, and superintendents. Significant differences in the variable perception composite

score were measured with a one-way analysis of variance (ANOVA) to answer Research Questions 2 through 5. This measurement was chosen due to the use of the single dependent variable and independent variables with three or more levels. Research Question 6 utilized multiple regression analysis. This measurement was chosen to see if the predictor variables could predict any variance from the criterion variable.

The independent variables of this study consisted of educational position type, the IDOE A – F Accountability designation for the 2012-2013 school year, the school or district socioeconomic status (SES) level, and the location of the school or district. The educational position type independent variable had three levels: teacher, principal, or superintendent. The independent variable of location had three levels: suburban, urban, or rural. The independent variable based on SES level had three levels: 0%–33%, 34%–67%, or 68%–100%. The IDOE A–F Accountability designation for the 2012-2013 school year, as an independent variable, had four levels: A, B, C, D or F. The dependent variables were the variable perception survey composite scores as they related to educational position type, location, SES level, and letter grade received from the IDOE during the 2012–13 school year.

For the final research question, the composite score served as the criterion variable. For multiple regression purposes, there were five different predictor variables. These predictor variables were specific SES level, school enrollment size, building specific accountability grade, location type, and the number of years the principal had been in his or her current position. By using this specific data, the predictor variables gave me a set of scale scores to run the multiple regression analysis.

The testing of assumptions (homogeneity of variance, normality, and independence) were tested before the ANOVA was conducted. A Tukey's HSD post hoc test was required if a

significant one-way ANOVA result occurred to determine which groups were significantly different from one another. If a violation of the assumption of homogeneity of variance occurred while testing the assumptions, a Games-Howell post hoc test was utilized instead of the Tukey HSD because this test does not require equal variances. Finally, I checked whether any of the dependent variable scores within the groups was an outlier. If so, these scores were removed from the ANOVA model prior to running the test.

CHAPTER 4

ANALYSES OF DATA

The main objective of this quantitative study was to identify teacher, principal, and superintendent perceptions on the variables that are used to measure school accountability. Recent research suggests that reliability issues exist due to the lack of consistency when utilizing AYP to measure school accountability (Elmore, 2002; Guilfoyle, 2006; Linn & Haug, 2002; Maleyko & Gawlik, 2011; Porter et al., 2005; Schoen & Fusarelli, 2008; Staiger et al., 2002; Wiley et al., 2005). Linn (2008) argued that a state's ability to identify a successful school or a failing school may be inconsistent due to the methodology that is utilized by the state to measure AYP rather than the actual effectiveness of schools. This study focused on the variables that are used to measure school accountability and their significance as indicators of school performance. "The No Child Left Behind Act of 2001 requires states to establish an accountability system that is both valid and reliable" (NCLB, 2002). As evidenced by ongoing policy debates, the question of what, who and how to measure school performance continues to plague the validity of NCLB (Gong et al., 2006).

For the purposes of this study, a survey methodology was utilized to collect data from teachers, principals, and superintendents who worked in the public school setting in the state of Indiana. Participants were asked about their perceptions on common variables used to measure school accountability. In addition, teachers, principals, and superintendents also had the

opportunity to express alternative variables they believed would better measure schools, as well as offer suggestions to improve Indiana's accountability system.

In an attempt to quantitatively measure the perceptions on the variables used to measure school accountability, I created the Perceptions of Variables Utilized to Effectively Measure School Accountability Survey. The review of literature indicated that numerous qualitative-based research projects have been conducted on the concept of school accountability, but little evidence exists of survey-based research that explores the perceptions of different job roles and their beliefs on what they feel should be utilized to measure school accountability.

The sampling procedure Chapter 3 was followed during distribution of The Perceptions of Variables Utilized to Effectively Measure School Accountability Survey. I attempted to e-mail the informed consent and the survey link to every public school teacher, principal, and superintendent in the state of Indiana. I was able to procure the e-mail addresses of teachers, principals, and superintendents through a freedom of information request via the IDOE. There were 746 participants that chose to complete the Qualtrics designed survey.

Descriptive Analyses

Whole Sample Demographic Data

The survey process sought out demographic data and utilized SPSS software to report the whole sample's current position type, percentage of students who received free or reduced lunches, the Indiana A–F accountability grade, and the location type of their school/district. A total of 746 individuals participated in the study. The current position of the respondents included 471 (63.1%) teachers, 209 (28.0%) principals, 59 (7.9%) superintendents, and seven respondents (.9%) who did not identify their position type.

Participants revealed the SES of their school by indicating the free/reduced lunch rates.

For teachers and principals, 344 participants (46.1%) noted 34%–67% free/reduced lunch rate, 214 (28.7%) noted between 68%–100% free/reduced lunch rate, 114 (15.3%) participants signified a free/reduced lunch rate less than 33%, and 74 (9.9%) did not identify their percentage. At the district level, 38 participants (64.4%) noted 34%–67% free/reduced lunch rate, 18 (30.5%) signified a free/reduced lunch rate to be less than 33%, and three (5.0%) identified between 68%–100%.

For teachers and principals, 280 (37.5%) participants identified their school accountability grade as an A grade, 155 (20.8%) identified B grade, 112 (15.0%) identified C grade, 115 (15.4%) identified D/F grade, and 84 (11.3%) did not indicate a school grade. At the district level, 29 (49.1%) participants identified their district accountability grade as an A grade, 16 (27.1%) identified B grade, 11 (18.6%) identified C grade, two (3.3%) identified D/F grade, and one (1.6%) did not indicate a district grade.

The final demographic indicator required respondents to identify their school or district geographic locations by choosing between urban, rural, and suburban. Three hundred eleven (41.7%) teachers and principals classified their schools as rural, 230 (30.8%) classified urban, 187 (25.1%) classified suburban, and 18 (2.4%) did not indicate their geographic locations. At the district level, 41 (69.5%) participants classified their districts as rural, 14 (23.7%) classified suburban, three (5.0%) classified urban, and one (1.6%) did not indicate his or her geographic location.

Whole Sample Descriptive Data–Variable Belief

The 746 respondents were asked questions related to their belief in a specific accountability statement. These statements included distinctive variables that were utilized in the current Indiana accountability system or variables used in other common accountability

models. Within the sample, all 746 participants answered 10 questions related to the use of specific variables to measure school accountability. Six choices were available for participants to select from when responding to each question. The choices were 1 = *strongly disagree*, 2 = *disagree*, 3 = *somewhat disagree*, 4 = *somewhat agree*, 5 = *agree*, and 6 = *strongly agree*.

An analysis of the participant responses from the whole sample data identified which of the variables received the highest mean scores. The top three whole-sample responses were the inclusion of a student's SES into the accountability model, the ability to gain bonus points for showing positive growth, and observing individual student growth from one year to the next. The average responses from the whole sample can be found in Table 2.

Further analysis of the participant responses identified the variables that received the lowest mean scores. The whole sample responses showed that including student gender into an accountability model was not favored. In terms of how student performance data were utilized, there was not the strong belief in an accountability system that penalizes a school for low growth. Finally, the whole sample did not favor having a specific yearly target, similar to AYP, that students were expected to meet.

Table 2

Whole Sample Data for Variable Belief Statement

Variable	*M*	*SD*
Use of ISTEP+ or ECA	3.12	1.39
Growth of bottom 25%	3.68	1.35
Separation of top 75%	3.46	1.30
Penalization for low growth	2.34	1.22
Awarding of bonus for high growth	4.16	1.39

Table 2 (Continued)

Variable	*M*	*SD*
Specific percentage every year	2.50	1.32
Individual student growth over time	4.06	1.28
Inclusion of teacher effectiveness	3.68	1.31
Inclusion of SES status	4.73	1.33
Inclusion of student race	2.60	1.41
Inclusion of student gender	2.17	1.23

Demographic Statistics by Position Type–Teacher

A total of 471 teachers participated in the study. Participants revealed the SES of their school by indicating the free/reduced lunch rate. Two hundred twenty-nine (48.6%) participants noted a 34%–67% free/reduced lunch rate, 170 (36.1%) participants noted between 68%–100% free/reduced lunch rate, 67 (14.2%) participants signified a free/reduced lunch rate to be less than 33%, and five (1.1%) participants did not identify their percentage.

One hundred eighty-seven (39.5%) participants identified their school accountability grade as an A grade, 103 (21.9%) identified B grade, 81 (17.2%) identified C grade, 87 (18.5%) identified D/F grade, and 13 (2.8%) did not indicate a school grade.

The final demographic indicator required respondents to identify their school geographic locations by choosing between urban, rural, and suburban. One hundred seventy-six (37.4%) teachers classified their school as rural, 172 (36.5%) classified urban, 117 (24.8%) classified suburban, and six (1.3%) did not indicate their geographic locations.

Descriptive Statistics by Position Type–Teacher

Each teacher participant was also asked questions related to their belief in a specific accountability statement. These statements included distinctive variables that are utilized in the current Indiana accountability system or variables used in other common accountability models.

An analysis of the participant responses from the teacher sample data identified which of the variables received the highest mean scores. The top three teacher sample responses were the inclusion of a student's SES into the accountability model, the ability to gain bonus points for showing positive growth, and, observing individual student growth from one year to the next. The average responses from the teacher sample can be found in Table 3. When comparing teacher sample data to the whole sample group, the findings were the same. The highest mean average for both the whole sample and the teacher sample was the variable that included socioeconomic information into the accountability model.

Further analysis of the participant responses identified the variables that received the lowest mean scores. The teacher sample response showed that including student gender into an accountability model was not favored. This belief mirrored the whole sample response. In terms of how student performance data was utilized, both the whole sample and the teacher sample data found there was not a strong belief in an accountability system that penalized a school for low growth. Finally, the teacher sample did not favor having a specific yearly target, similar to AYP, that students were expected to meet.

Table 3

Teacher Sample Data for Variable Belief Statement

Variable	*M*	*SD*
Use of ISTEP+ or ECA	2.78	1.31
Growth of bottom 25%	3.53	1.37
Separation of top 75%	3.38	1.30
Penalization for low growth	2.22	1.16
Awarding of bonus for high growth	4.07	1.44
Specific percentage every year	2.27	1.23
Individual student growth over time	3.97	1.35
Inclusion of teacher effectiveness	3.56	1.36
Inclusion of SES status	4.94	1.23
Inclusion of student race	2.60	1.41
Inclusion of student gender	2.14	1.21

Because the teacher sample responses had similar variables that received the highest and lowest mean score, continued analysis of the data was conducted to identify the variables that had the largest mean score change from the whole sample. The largest change occurred with the teacher sample response for the utilization of ISTEP+ or ECA scores where there was a .34 decrease in the mean average over the whole sample response for the same variable. A similar decrease occurred with the teacher sample response for the variable of a specific yearly target, similar to AYP. The teacher response sample mean average had a .23 decrease from the whole sample mean score. There was, however, an increase in the teacher sample mean score for

including the SES into an accountability model. The teacher sample mean score was .21 higher than the whole group sample for the same variable.

Demographic Statistics by Position Type–Principal

A total of 209 principals participated in the study. Participants revealed the SES of their schools by indicating the free/reduced lunch rate. One hundred fifteen (55.0%) participants noted a 34%–67% free/reduced lunch rate, 44 (21.1%) noted between 68%–100% free/reduced lunch rate, 47 (22.5%) participants signified a free/reduced lunch rate to be less than 33%, and three (1.4%) did not identify their percentage.

Ninety-three (44.5%) participants identified their school accountability grade was an A grade, 52 (24.9%) identified B grade, 31 (14.8%) identified C grade, 28 (13.4%) identified D/F grade, and five (2.4%) did not indicate a school grade.

The final demographic indicator required that respondents identify their school geographic location by choosing between urban, rural, and suburban. Ninety-five (45.5%) principals classified their schools as rural, 56 (26.8%) classified suburban, 55 (24.8%) classified urban, and three (1.4%) did not indicate their geographic locations.

Descriptive Statistics by Position Type–Principal

Each principal participant was asked questions related to their belief in a specific accountability statement. These statements included distinctive variables that were utilized in the current Indiana accountability system or variables used in other common accountability models.

An analysis of the participant responses from the principal sample data identified which of the variables received the highest mean scores. According to principal participants, the variable of awarding bonus points for high growth in an accountability model attained the

highest mean score. Along with the teacher sample, the principal respondents expressed that a student's SES should be included in an accountability model. Finally, the principal sample response indicated the support of observing individual student growth from year to year. The average responses from the principal sample can be found in Table 4. Similar to the teacher sample data, the principal data mirrored the whole group data in regard to the variables that had the highest mean scores.

Table 4

Principal Sample Data for Variable Belief Statement

Variable	*M*	*SD*
Use of ISTEP+ or ECA	3.67	1.29
Growth of bottom 25%	3.92	1.24
Separation of top 75%	3.59	1.26
Penalization for low growth	2.61	1.26
Awarding of bonus for high growth	4.47	1.19
Specific percentage every year	2.95	1.36
Individual student growth over time	4.26	1.06
Inclusion of teacher effectiveness	3.84	1.17
Inclusion of SES status	4.36	1.40
Inclusion of student race	2.69	1.43
Inclusion of student gender	2.27	1.28

Further analysis of the participant responses identified the variables that received the lowest mean scores. Principals indicated the lowest level of agreement with the same variables as the teacher sample group. Data showed that principals did not favor the use of student gender in an accountability model. In terms of how student performance data were utilized, principal respondents did not support an accountability system that penalized a school for low growth. Finally, the practice of measuring school accountability similar to AYP received a low mean score from participants within the principal sample. When comparing the principal mean data to the whole group mean data in Table 2, the variables with the lowest belief averages were the same.

Since the principal sample responses had similar variables that received the highest and lowest mean scores, continued analysis of the data were conducted to identify the variables that had the largest mean score change from the whole sample. The largest change occurred with the principal sample response for the use of ISTEP+ or ECA scores to measure school accountability, where the principal mean score increased .55 when compared to the whole sample response for the same variable. A similar increase occurred with the principal sample response for the variable of a specific yearly target, similar to AYP. The principal sample mean average increased .45 from the whole sample mean score. A decrease occurred in the principal sample mean score for including the SES into an accountability model. The principal sample mean score decreased .37 when compared to the whole group sample for the same variable.

Demographic Statistics by Position Type–Superintendent

A total of 59 superintendents participated in the study. Participants revealed the SES of their districts by indicating the free/reduced lunch rate. Thirty-eight (64.4%) participants noted a 34%–67% free/reduced lunch rate, three (5.1%) noted between 68%–100% free/reduced lunch

rate, and 18 (30.5%) participants signified a free/reduced lunch rate to be less than 33%.

Twenty-nine (49.2%) participants identified their district accountability grade as an A grade, 16 (27.1%) identified a B grade, 11 (18.6%) identified a C grade, two (3.4%) D/F grade, and one (1.7%) did not indicate a district grade.

The final demographic indicator required respondents to identify their district geographic location by choosing between urban, rural, and suburban. Forty (67.8%) superintendents classified their districts as rural, 14 (23.7%) classified suburban, three (5.1%) classified urban, and one (1.7%) did not indicate his or her geographic location.

Descriptive Statistics by Position Type–Superintendent

Each superintendent participant was also asked questions related to their belief in a specific accountability statement. These statements included distinctive variables that are utilized in the current Indiana accountability system or variables used in other common accountability models.

An analysis of the participant responses from the superintendent sample data identified which of the variables received the highest mean scores. Responses from the sample group indicated that superintendents favored taking the SES of students into account when determining school accountability. This specific variable consistently received the highest mean scores among all four sampling groups. The superintendent sample, along with the teacher and principal sample, identified that an accountability model should look at the growth of student data over time. Finally, responses from the superintendent sample data also identified that including teacher effectiveness into an accountability model would be favored. Respondents from this specific sample were the only group that ranked an accountability model utilizing

teacher effectiveness among the top three variables. The average responses from the principal sample can be found in Table 5.

Table 5

Superintendent Sample Data for Variable Belief Statement

Variable	*M*	*SD*
Use of ISTEP+ or ECA	3.95	1.40
Growth of bottom 25%	4.02	1.33
Separation of top 75%	3.63	1.33
Penalization for low growth	2.41	1.42
Awarding of bonus for high growth	3.98	1.43
Specific percentage every year	2.81	1.38
Individual student growth over time	4.27	1.19
Inclusion of teacher effectiveness	4.15	1.22
Inclusion of SES status	4.34	1.46
Inclusion of student race	2.41	1.43
Inclusion of student gender	2.07	1.26

Further analysis of the participant responses identified the variables that received the lowest mean scores. Data showed that including student gender and race into an accountability model was not favored. Participants from the four sample groups consistently scored the variable of gender as the lowest. Similarly, respondents from all four sample groups, expressed that an accountability system which penalizes an organization for low growth should not be utilized.

Because the superintendent sample responses had similar variables that received the highest and lowest mean scores, continued analysis of the participant data identified variables with the largest mean score change from the whole sample. The largest change occurred with the superintendent sample response for the use of ISTEP+ or ECA scores to measure school accountability, where the mean average increased .83 from the whole sample response for the same variable. A similar increase occurred with the superintendent sample response in regard to including teacher effectiveness into an accountability model. The superintendent sample mean average increased .47 from the whole sample mean score. A decrease occurred in the superintendent mean score when an accountability model included student race. The superintendent sample mean score decreased .19 when compared to the whole group sample for the same variable.

Demographic Statistics by Indiana A–F Accountability Grade–A

Three hundred nine participants identified their school/district as an A school. The current position of the respondents included 187 (66.8%) teachers, 93 (33.2%) principals, and 29 (9.3%) superintendents. The response rates for the A sample were consistent with the whole sample distribution of 63% teachers, 28% principals, and 9% superintendents.

Participants revealed the SES of their school by indicating the free/reduced lunch rate. For teachers and principals, 146 (52.1%) noted 34%–67% free/reduced lunch rate, 93 (33.2%) noted between 68%–100% free/reduced lunch rate, and 41 (14.6%) participants signified a free/reduced lunch rate less than 33%. At the district level, 15 superintendents (51.7%) noted a 34%–67% free/reduced lunch rate, 13 (44.8%) signified a free/reduced lunch rate less than 33%, and one (3.4%) identified a free/reduced lunch rate between 68%–100%.

The final demographic indicator required respondents to identify their schools' or districts' geographic locations by choosing between urban, rural, and suburban. One hundred thirty-five (48.2%) teachers and principals classified their schools as rural, 106 (37.9%) classified urban, and 38 (13.6%) classified suburban. At the district level, 18 (62.1%) participants classified their districts as rural, 11 (37.9%) classified suburban, and zero (0%) as urban.

Descriptive Statistics by Indiana A–F Accountability Grade–A

Each participant was asked questions related to his or her belief in a specific accountability statement. These statements included distinctive variables that are utilized in the current Indiana accountability system or variables used in other common accountability models.

An analysis of the responses from the A school sample data identified the variables that received the highest mean scores from teachers and principals. The responses with the highest mean average were: utilizing the student SES, awarding bonus points for high growth, and utilizing a growth model to observe student performance. When comparing the A school location sample data to the whole sample data, the same variables received the highest mean average. The average response data from the A school sample data can be found in Table 6.

At the district level, an analysis of the A district sample data identified the variables that received the highest mean score from superintendents. The responses with the highest mean average were: including teacher effectiveness into an accountability model ($M = 4.52$, $SD = 1.02$), the use of a growth style model ($M = 4.48$, $SD = 1.06$), and the utilization of ISTEP+ or ECA to measure school accountability ($M = 4.24$, $SD = 1.24$). Compared to both the school level A school sample data, as well as the whole sample data, the superintendent responses differed. Agreement among participants existed with the use of a growth-style model; however, this was

the only sample group that advocated to utilize ISTEP+ or ECA data, as well as to embed teacher effectiveness into the accountability model.

Table 6

Indiana A–F Accountability Letter Grade Sample Data for Variable Belief Statement–A School

Variable	*M*	*SD*
Use of ISTEP+ or ECA	3.14	1.40
Growth of bottom 25%	3.64	1.34
Separation of top 75%	3.41	1.32
Penalization for low growth	2.40	1.19
Awarding of bonus for high growth	4.25	1.35
Specific percentage every year	2.55	1.34
Individual student growth over time	4.02	1.30
Inclusion of teacher effectiveness	3.71	1.29
Inclusion of SES status	4.67	1.28
Inclusion of student race	2.58	1.35
Inclusion of student gender	2.06	1.07

Further analysis of the participant responses identified the variables that received the lowest mean scores. Data showed that including student gender into an accountability model was not favored. Utilizing student gender received the lowest mean score among the whole population. Similarly, an accountability system that penalizes a school/district due to low growth received low mean scores from respondents in both the A school sample and the whole sample. Finally, teachers and principals in the A school sample expressed that a system similar

to AYP should not be utilized. At the district level, superintendents denoted that gender (M = 1.93, SD = 1.22) should not be a variable considered when measuring school accountability. Superintendents also did not favor utilizing a system that penalized districts for low growth (M =2.48, SD = 1.48).

Because the A school sample responses had similar variables that received the highest and lowest mean scores, continued analysis of the participant data identified variables with the largest mean score change from the whole sample. The largest change occurred with the A sample response for the inclusion of gender into a school accountability model, where a .11 decrease occurred in the mean average from the whole sample response for the same variable. An increase occurred with the A school sample response for the variable of awarding bonus points for high growth. The A school sample mean average had a .09 increase from the whole sample mean score.

Demographic Statistics by Indiana A–F Accountability Grade–B

A total of 171 participants identified their school/districts received a B accountability grade. The current position of the respondents included 103 (60.2%) teachers, 52 (30.4%) principals, and 16 (9.4%) superintendents. The response rates for the B sample were consistent with the whole sample distribution of 63% teachers, 28% principals, and 9% superintendents.

Participants revealed the SES of their school by indicating the free/reduced lunch rate. For teachers and principals, 94 (60.6%) noted a 34%–67% free/reduced lunch rate, 45 (29.0%) between 68%–100% free/reduced lunch rate, and 16 (10.3%) participants signified a free/reduced lunch rate less than 33%. At the district level, 11 superintendents (68.8%) noted a 34%–67% free/reduced lunch rate, four (25.0%) signified a free/reduced lunch rate less than 33%, and one (6.3%) identified between 68%–100%.

The final demographic indicator required respondents to identify their school or district geographic locations by choosing between urban, rural, and suburban. Seventy-one (45.8%) teachers and principals classified their school as rural, 50 (32.3%) classified urban, and 33 (21.3%) classified suburban. At the district level, 13 (81.3%) participants classified their districts as rural, two (12.5%) classified suburban, and one (6.3%) identified their district as urban.

Descriptive Statistics by Indiana A–F Accountability Grade–B

Participants were asked questions related to their belief in a specific accountability statement. These statements included distinctive variables that are utilized in the current Indiana accountability system or variables used in other common accountability models.

An analysis of the responses from the B school sample data identified the variables that received the highest mean scores from teachers and principals. The responses with the highest mean average were: utilizing the student SES, awarding bonus points for high growth, and utilizing a growth model to observe student performance. When comparing the B school sample data to the whole sample data, participant mean scores were the highest for the same variables. The average response data from the B school sample data can be found in Table 7.

At the district level, an analysis of the B district sample data identified the variables that received the highest mean score from superintendents. The responses with the highest mean average were: including SES into an accountability model ($M = 4.69$, $SD = 1.49$), looking at the growth of the bottom 25% ($M = 4.19$, $SD = 1.52$), and the use of a growth style model ($M = 3.94$, $SD = 1.39$). Compared to both the school level B school sample data, as well as the whole sample data, the superintendent responses differed. There was agreement with the use of a growth-style model and that student performance should be assessed over multiple years;

however, this was the only sample group that advocated to assess growth of the bottom 25% of students.

Table 7

Indiana A–F Accountability Letter Grade Sample Data for Variable Belief Statement–B School

Variable	*M*	*SD*
Use of ISTEP+ or ECA	3.03	1.36
Growth of bottom 25%	3.62	1.32
Separation of top 75%	3.45	1.33
Penalization for low growth	2.33	1.21
Awarding of bonus for high growth	4.13	1.42
Specific percentage every year	2.37	1.23
Individual student growth over time	3.92	1.33
Inclusion of teacher effectiveness	3.43	1.36
Inclusion of SES status	4.69	1.36
Inclusion of student race	2.57	1.42
Inclusion of student gender	2.21	1.18

Further analysis of the participant responses identified the variables that received the lowest mean scores. Data showed that including student gender into an accountability model was not favored. Utilizing student gender received the lowest mean score among the whole population. Similarly, an accountability system that penalizes a school/district due to low growth received low mean scores from respondents in both the B school sample and the whole sample. Finally, teachers and principals in the B school sample expressed that a system similar

to AYP should not be utilized. At the district level, superintendents noted that gender ($M = 2.50$, $SD = 1.55$) should not be a variable considered when measuring school accountability. Superintendents also did not favor utilizing a system that penalized districts for low growth ($M = 2.56$, $SD = 1.59$).

Because the B school and district sample responses had similar variables that received the highest and lowest mean scores, continued analysis of the participant data identified variables with the largest mean score change from the whole sample. The largest mean difference occurred with the B sample response for the inclusion of teacher effectiveness into a school accountability model, evidenced by a .25 decrease in the mean average over the whole sample response for the same variable. An increase occurred with the B district sample response for observing the growth of the bottom 25%. The B district sample mean difference scored .51 higher than the whole sample mean score.

Demographic Statistics by Indiana A–F Accountability Grade–C

A total of 171 participants identified their school/districts received a C accountability grade. The current position of the respondents included 81 (65.9%) teachers, 31 (25.2%) principals, and 11 (8.9%) superintendents. The response rates for the C samples were consistent with the whole sample distribution of 63% teachers, 28% principals, and 9% superintendents.

Participants revealed the SES of their schools by indicating the free/reduced lunch rate. For teachers and principals, 63 (65.9%) noted a 34%–67% free/reduced lunch rate, 43 (38.4%) noted between 68%–100% free/reduced lunch rate, and five (4.5%) participants signified a free/reduced lunch rate to be less than 33%. At the district level, all 11 superintendents (100%) noted a 34%–67% free/reduced lunch rate.

The final demographic indicator required respondents to identify their school or district geographic locations by choosing between urban, rural, and suburban. Fifty-two (46.4%) teachers and principals classified their schools as urban, 43 (38.4%) classified urban, and 16 (14.3%) classified suburban. At the district level, nine (81.8%) participants classified their districts as rural, one (9.1%) chose suburban, and one (9.1%) identified the district as urban.

Descriptive Statistics by Indiana A–F Accountability Grade–C

Participants were asked questions related to their belief in a specific accountability statement. These statements included distinctive variables that are utilized in the current Indiana accountability system or variables used in other common accountability models.

An analysis of the responses from the C school sample data identified the variables that received the highest mean scores from teachers and principals. The responses with the highest mean average were: utilizing the student SES, awarding bonus points for high growth, and utilizing a growth model to observe student performance. The same variables received the highest mean average, when comparing the C school sample data to the whole sample data. The average response data from the C school sample can be found in Table 8.

At the district level, an analysis of the C district sample data identified the variables that received the highest mean score from superintendents. The responses with the highest mean average were: including SES into an accountability model ($M = 4.82$, $SD = 1.47$), rewarding schools for high growth ($M = 4.18$, $SD = 1.25$), and the use of a growth style model ($M = 4.09$, $SD = 1.22$). Compared to both the C school sample data, as well as the whole sample data, the superintendent responses did not differ.

Table 8

Indiana A–F Accountability Letter Grade Sample Data for Variable Belief Statement–C School

Variable	*M*	*SD*
Use of ISTEP+ or ECA	3.08	1.36
Growth of bottom 25%	3.62	1.36
Separation of top 75%	3.55	1.21
Penalization for low growth	2.41	1.24
Awarding of bonus for high growth	4.15	1.40
Specific percentage every year	2.45	1.26
Individual student growth over time	4.14	1.21
Inclusion of teacher effectiveness	3.63	1.36
Inclusion of SES status	4.83	1.24
Inclusion of student race	2.78	1.50
Inclusion of student gender	2.18	1.27

Further analysis of the participant responses identified the variables that received the lowest mean scores. When comparing the C school sample responses and the whole sample distribution, data suggested that the inclusion of student gender into an accountability model was not favored. Utilizing student gender received the lowest mean score among the whole population. Similarly, an accountability system that penalizes a school/district due to low growth received low mean scores from respondents in both the C school sample and the whole sample. Finally, teachers and principals in the C school sample expressed that a system similar to AYP should not be utilized. At the district level, superintendents noted that gender ($M = 1.73$,

SD = .647) and race (*M* = 1.91, *SD* = .701) should not be variables considered when measuring school accountability. Superintendents also did not favor utilizing a system that penalized districts for low growth (*M* = 2.00, *SD* = 1.00). The only difference in responses from the "C" district sample was that superintendents did not support utilizing race as a variable when compared to the whole sample responses.

Because the C school and district sample responses had similar variables that received the highest and lowest mean scores, continued analysis of the participant data identified variables with the largest mean score difference from the whole sample. The "C" school sample mean average increased .18 from the whole sample mean score when considering student race as an accountability model variable. When compared to the whole sample mean score, a .69 decrease occurred with the "C" district sample response for including student race.

Demographic Statistics by Indiana A–F Accountability Grade–D/F

A total of 117 participants identified their school/districts received D/F accountability grade. The current position of the respondents included 87 (74.4%) teachers, 28 (23.9%) principals, and two (1.7%) superintendents. The response rates for the D/F schools differed from the whole sample distribution. The D/F school sample showed an 11% increase in teacher respondents, as well as a 7% decrease in superintendent responses.

Participants revealed the SES of their schools by indicating the free/reduced lunch rate. For teachers and principals, 79 (68.7%) participants noted 68%–100% free/reduced lunch rate, 35 (30.4%) noted between 34%–67% free/reduced lunch rate, and zero (0%) participants signified a free/reduced lunch rate to be less than 33%. At the district level, one superintendent (50.0%) noted a 34%–67% free/reduced lunch rate and one (50%) signified a free/reduced lunch rate between 68%–100%.

The final demographic indicator required respondents to identify their school or district geographic locations by choosing between urban, rural, and suburban. Eighty-one (70.4%) teachers and principals classified their school as urban, 18 (15.7%) classified rural, and 15 (13.0%) classified suburban. At the district level, one (50.0%) participant classified his or her district as rural and one (50%) identified his or her district as urban.

Descriptive Statistics by Indiana A–F Accountability Grade–D/F

Each participant was asked questions related to his or her belief in a specific accountability statement. These statements included distinctive variables that are utilized in the current Indiana accountability system or variables used in other common accountability models.

An analysis of the responses from the D/F school sample data identified the variables that received the highest mean scores from teachers and principals. The responses attaining the highest mean average were: utilizing the SES status, awarding bonus points for high growth, and utilizing a growth model to observe student performance. The same variables received the highest mean average scores, when comparing the D/F school sample data to the whole sample data. The average response data from the D/F school type can be found in Table 9.

At the district level, an analysis of the D/F district sample data identified the variables that received the highest mean score from superintendents. Superintendents believed that teacher effectiveness ($M = 5.00$, $SD = .00$), a growth style model ($M = 5.00$, $SD = 1.41$), and rewarding schools for high growth ($M = 4.00$, $SD = 1.41$) should be included into an accountability model. When comparing both the D/F school sample data, as well as the whole sample data, the superintendent responses differed slightly. Superintendent responses suggested utilizing teacher effectiveness where the other samples did not.

Table 9

Indiana A–F Accountability Letter Grade Sample Data for Variable Belief Statement–D/F School

Variable	*M*	*SD*
Use of ISTEP+ or ECA	2.92	1.29
Growth of bottom 25%	3.74	1.46
Separation of top 75%	3.50	1.25
Penalization for low growth	2.14	1.23
Awarding of bonus for high growth	4.26	1.41
Specific percentage every year	2.58	1.37
Individual student growth over time	4.32	1.19
Inclusion of teacher effectiveness	3.85	1.25
Inclusion of SES status	4.94	1.42
Inclusion of student race	2.67	1.50
Inclusion of student gender	2.41	1.58

Further analysis of the participant responses identified the variables that received the lowest mean scores. Data showed that including student gender into an accountability model was not favored. Utilizing student gender received the lowest mean score among each sample population. Similarly, an accountability system that penalizes a school/district due to low growth received low mean scores from respondents in both the D/F school sample and the whole sample. Finally, teachers and principals in the D/F school sample expressed that a system similar to AYP should not be utilized. At the district level, superintendents were in agreement

that gender (M = 1.50, SD = .707) and race (M = 1.50, SD = .707) should not be a variable considered when measuring school accountability. Superintendents also did not favor utilizing a system that penalized districts for low growth (M = 1.50, SD = .707). The D/F district sample's only difference was that superintendents did not support utilizing race as a variable when compared to the D/F school sample and the total sample.

Because the D/F school and district sample responses had similar variables that received the highest and lowest mean scores, continued analysis of the participant data identified variables with the largest mean score change from the whole sample. The largest mean difference occurred with the D/F school sample response for an accountability model that observed student growth over time, increasing .26 in the mean average over the whole sample response for the same variable. This was followed closely by a mean difference of .24 in which the D/F school sample had a higher mean average for including student gender over the whole sample response. With only two superintendents responding in this category, the mean difference was not observed.

Demographic Statistics by SES: 0%–33% Free/Reduced Lunch

A total of 131 participants identified their school/districts had a socioeconomic make-up of 0%–33%. The current position of the respondents included 67 (51.1%) teachers, 47 (35.9%) principals, and 17 (13.0%) superintendents. The response rate was higher for both principals and superintendents when compared to the whole sample distribution.

For teachers and principals, 93 (81.6%) participants identified their school accountability grade as A grade, 16 (14.0%) identified B grade, five (4.4%) identified C grade, and zero (0%) received a D/F grade. At the district level, 13 (72.2%) participants identified their district accountability grades as A grade and four (22.2%) identified B grade. None of the

superintendents received C, D or F as letter grades.

The final demographic indicator required respondents to identify their school or district geographic location by choosing between urban, rural, and suburban. Sixty-one (53.5%) teachers and principals classified their school as suburban, 50 (43.9%) identified rural, and two (1.8%) identified suburban. At the district level, 11 (61.1%) participants classified their districts as rural and six (33.3%) identified their districts as suburban.

Descriptive Statistics by SES: 0%–33% Free/Reduced Lunch

Participants were asked questions related to their beliefs in a specific accountability statement. These statements included distinctive variables that are utilized in the current Indiana accountability system or variables used in other common accountability models.

An analysis of the responses from the 0%–33% school sample data identified the variables that received the highest mean scores from teachers and principals. The responses with the highest mean average were: utilizing the student SES, awarding bonus points for high growth, and utilizing a growth model to observe student performance. The same variables received the highest mean scores, when comparing the 0%–33% school sample data to the whole sample data. The average response data from the 0%–33% school type can be found in Table 10.

At the district level, an analysis of the 0%–33% district sample data identified the variables that received the highest mean score from superintendents. The responses with the highest mean average were: taking the SES into account ($M = 4.17$, $SD = 1.34$), observing the growth of the bottom 25% of students ($M = 4.00$, $SD = 1.24$), and embedding teacher effectiveness into an accountability model ($M = 3.94$, $SD = 1.26$). Compared to the 0%–33% school sample data, as well as the whole sample data, the superintendent responses were

different. Superintendents were the only ones who felt that teacher effectiveness, as well as the bottom 25% should be included in an accountability model.

Table 10

Variable Belief Statement by School SES: 0%–33% Free/Reduced Lunch

Variable	*M*	*SD*
Use of ISTEP+ or ECA	3.25	1.39
Growth of bottom 25%	3.66	1.40
Separation of top 75%	3.41	1.35
Penalization for low growth	2.39	1.17
Awarding of bonus for high growth	4.35	1.27
Specific percentage every year	2.67	1.41
Individual student growth over time	4.04	1.29
Inclusion of teacher effectiveness	3.73	1.30
Inclusion of SES status	4.49	1.33
Inclusion of student race	2.46	1.32
Inclusion of student gender	2.09	1.16

Further analysis of the participant responses identified the variables that received the lowest mean scores. At the school level, data showed that including student gender into an accountability model was not favored. Utilizing student gender received the lowest mean score among the respondents from each sample population. Similarly, an accountability system that penalizes a school/district due to low growth received low mean scores from respondents in both the 0%–33% school sample and the whole sample. Finally, teachers and principals in the 0%–

33% school sample expressed that a system similar to AYP should not be utilized. At the district level, superintendents were in agreement that gender (M = 2.28, SD = 1.27) and race (M = 2.56, SD = 1.42) should not be a variable considered when measuring school accountability. Superintendents also did not favor utilizing a system that penalized districts for low growth (M = 2.11, SD = 1.49). Again, there were no differences among the school, district, or total sample.

Because the 0%–33% school and district sample responses had similar variables that received the highest and lowest mean scores, continued analysis of the participant data identified variables with the largest mean score change from the whole sample. The largest mean difference occurred with the 0%–33% school sample response for the inclusion of SES into a school accountability model, evidenced by a .25 decrease in the mean average over the whole sample response for the same variable. The 0%–33% school sample increased .19 in the mean average over the whole sample for awarding bonus points for high growth. A decrease occurred with the 0%–33% district sample response for including SES. The 0%–33% district sample mean score decreased .56 from the whole sample mean score.

Demographic Statistics by SES: 34%–67% Free/Reduced Lunch

A total of 382 participants identified their school/districts had a socioeconomic make-up of 34%–67%. The current position of the respondents included 229 (59.9%) teachers, 115 (30.1%) principals, and 38 (9.9%) superintendents.

For teachers and principals, 146 (42.4%) participants identified their school accountability grades as A grade, 94 (27.3%) identified B grade, 63 (18.3%) identified C grade, and 35 (10.2%) received a D/F grade. At the district level, 15 (39.5%) participants identified their district accountability grade as A grade, 11 (28.9%) identified B grade, 11 (28.9%) identified C grade, and one (2.6%) identified D or F grade.

The final demographic indicator required respondents to identify their school or district geographic locations by choosing between urban, rural, and suburban. One hundred ninety-four (56.4%) teachers and principals classified their schools as rural, 77 (22.4%) classified suburban, and 70 (20.3%) classified urban. At the district level, 29 (76.3%) participants classified their districts as rural, seven (18.4%) classified suburban, and two (5.3%) identified their districts as urban.

Descriptive Statistics by SES: 34%–67% Free/Reduced Lunch

Each participant was also asked questions related to his or her belief in a specific accountability statement. These statements included distinctive variables that are utilized in the current Indiana accountability system or variables used in other common accountability models.

An analysis of the responses from the 34%–67% school sample data identified the variables that received the highest mean scores from teachers and principals. The responses with the highest mean average were: utilizing the student SES, awarding bonus points for high growth, and utilizing a growth model to observe student performance. The same variables received the highest mean scores, when comparing the 34%–67% school sample data to the whole sample data. The average response data from the 34%–67% school type can be found in Table 11.

At the district level, an analysis of the 34%–67% district sample data identified the variables that received the highest mean score from superintendents. The responses with the highest mean average were: utilizing a growth model to measure student progress ($M = 4.45$, $SD = 1.16$), taking the SES into account ($M = 4.39$, $SD = 1.48$), and embedding teacher effectiveness into an accountability model ($M = 4.18$, $SD = 1.21$). Compared to both the 34%–67% school sample data, as well as the whole sample data, the superintendent responses were slightly

different. Superintendents were the only group that scored the inclusion of teacher effectiveness at a high rate.

Table 11

Variable Belief Statement by School SES: 34%–67% Free/Reduced Lunch

Variable	*M*	*SD*
Use of ISTEP+ or ECA	3.12	1.37
Growth of bottom 25%	3.64	1.26
Separation of top 75%	3.40	1.28
Penalization for low growth	2.35	1.18
Awarding of bonus for high growth	4.08	1.42
Specific percentage every year	2.46	1.26
Individual student growth over time	4.02	1.24
Inclusion of teacher effectiveness	3.69	1.30
Inclusion of SES status	4.75	1.27
Inclusion of student race	2.64	1.42
Inclusion of student gender	2.17	1.20

Further analysis of the participant responses identified the variables that received the lowest mean scores. School data showed that including student gender into an accountability model was not favored. Utilizing student gender received the lowest mean score among each sample population. Similarly, an accountability system that penalizes a school/district due to low growth received low mean scores from respondents in both the 34%–67% school sample and the whole sample. Finally, teachers and principals in the 34%–67% school sample expressed

that a system similar to AYP should not be utilized. At the district level, superintendents were in agreement that gender (M = 2.28, SD = 1.27) and race (M = 2.56, SD = 1.42) should not be a variable considered when measuring school accountability. Superintendents also did not favor utilizing a system that penalized districts for low growth (M = 2.11, SD = 1.49). Again, there were no differences among the school, district or total samples.

Because the 34%–67% school and district sample responses had similar variables that received the highest and lowest mean scores, continued analysis of the participant data identified variables with the largest mean score change from the whole sample. The largest mean difference occurred with the 34%–67% school sample response for awarding bonus points for high growth, increasing .08 in the mean average over the whole sample response for the same variable. The 34%–67% school sample decreased .08 in the mean average over the whole sample for the utilizing a specific yearly target. An increase occurred with the 34%–67% district sample response for embedding teacher effectiveness into an accountability model. The 34%–67% district sample mean score increased .50 when compared to the whole sample mean score.

Demographic Statistics by SES: 68%–100% Free/Reduced Lunch

A total of 217 participants identified their school/district had a socioeconomic make-up of 68%–100%. The current position of the respondents included 170 (78.3%) teachers, 44 (20.3%) principals, and three (1.4%) superintendents. The response rate for teachers in this sample was much higher than the whole sample distribution. In addition, the response rate for superintendents was much lower than the whole sample distribution.

For teachers and principals, 41 (19.2%) participants identified their school accountability grade as A grades, 45 (21.0%) identified B grades, 43 (20.1%) identified C grades, and 79 (36.9%) received D/F grades. At the district level, one (33.3%) participants identified his or her

district accountability grade as an A grade, one (33.3%) identified a B grade, and one (33.3%) identified a D/F grade.

The final demographic indicator required respondents to identify their school or district geographic locations by choosing between urban, rural, and suburban. One hundred fifty-four (72.0%) teachers and principals classified their schools as urban, 34 (15.9%) classified rural, and 26 (12.1%) classified suburban. At the district level, one (33.3%) participants classified their district as rural, one (33.3%) classified suburban, and one (33.3%) identified urban.

Descriptive Statistics by SES: 68%–100% Free/Reduced Lunch

Participants were asked questions related to their belief in a specific accountability statement. These statements included distinctive variables that are utilized in the current Indiana accountability system or variables used in other common accountability models.

An analysis of the responses from the 68%–100% school sample data identified the variables that received the highest mean scores from teachers and principals. The responses with the highest mean average were: utilizing the student SES, awarding bonus points for high growth, and utilizing a growth model to observe student performance. The same variables received the highest mean scores, when comparing the 68%–100% school sample data to the whole sample data. The average response data from the 68%–100% school type can be found in Table 12.

At the district level, an analysis of the 68%–100% district sample data was conducted to determine which of the variables received the highest mean scores. The responses with the highest mean average were: embedding teacher effectiveness into an accountability model (M = 5.00, SD = 1.00), rewarding school for high growth (M = 5.00, SD = 1.00), and observing the growth of the bottom 25% of students (M = 5.00, SD = 1.00). Compared to both the 68%–100%

school sample data, as well as the whole sample data, the superintendent responses were different. Superintendents were the only participants who felt that teacher effectiveness, as well as the bottom 25%, should be included in an accountability model.

Table 12

Variable Belief Statement by School SES: 68%–100% Free/Reduced Lunch

Variable	*M*	*SD*
Use of ISTEP+ or ECA	2.84	1.34
Growth of bottom 25%	3.64	1.47
Separation of top 75%	3.55	1.29
Penalization for low growth	2.28	1.27
Awarding of bonus for high growth	4.30	1.36
Specific percentage every year	2.42	1.32
Individual student growth over time	4.15	1.32
Inclusion of teacher effectiveness	3.54	1.35
Inclusion of SES status	4.94	1.32
Inclusion of student race	2.69	1.45
Inclusion of student gender	2.24	1.31

Further analysis of the participant responses identified the variables that received the lowest mean scores. School data showed that including student gender in an accountability model was not favored. Utilizing student gender received the lowest mean score among each sample population. Similarly, an accountability system that penalizes a school/district due to low growth received low mean scores from respondents in both the 68%–100% school sample

and the whole sample. Finally, teachers and principals in the 68%–100% school sample expressed that a system similar to AYP should not be utilized. At the district level, superintendents were in agreement that gender (M = 2.33, SD = 1.53) and race (M = 3.00, SD = 2.65) should not be a variable considered when measuring school accountability. Superintendents also did not favor utilizing a system that penalized districts for low growth (M = 2.67, SD = .577) Again, there were no differences among the school, district or total sample.

Because the 68%–100% school and district sample responses had similar variables that received the highest and lowest mean scores, continued analysis of the participant data identified variables with the largest mean score change from the whole sample. The largest mean difference occurred with the 68%–100% school sample response for utilizing ISTEP+ or ECA scores to measure school accountability, where a .28 decrease occurred in the mean average over the whole sample response for the same variable. The 68%–100% school sample increased .14 in the mean average over the whole sample for awarding bonus when high growth is achieved. With only three superintendent responses, the mean difference was not analyzed.

Demographic Statistics by School/District Locality–Urban

A total of 230 individuals identified their school/districts as urban. The current position of the respondents included 172 (74.8%) teachers, 55 (23.9%) principals, and three (1.3%) superintendents. The teacher sample for urban schools was approximately 11% higher than the whole sample distribution. In addition, the superintendent response rate was almost 7% lower than the whole sample distribution.

Participants revealed the SES of their schools by indicating the free/reduced lunch rates. For teachers and principals, 154 participants (46.1%) noted a 68%–100% free/reduced lunch rates, 70 (30.4%) noted between 34%–67% free/reduced lunch rates, and two (0.9%) signified a

free/reduced lunch rates to be less than 33%. At the district level, two (66.7%) noted a 34%–67% free/reduced lunch rates and one (33.3%) participant signified a free/reduced lunch rate as being between 68%–100%.

For teachers and principals, 38 (16.5%) participants identified their school accountability grades as A grade, 50 (21.7%) identified B grade, 52 (22.6%) identified C grade, and 81 (35.2%) indicated a D/F grade. At the district level, one (33.3%) participant identified his or her district accountability grade as a B grade, one (33.3%) identified a C grade, and one (33.3%) indicated a D/F district grade.

Descriptive Statistics by School/District Locality–Urban

Participants were asked questions related to their belief in a specific accountability statement. These statements included distinctive variables that are utilized in the current Indiana accountability system or variables used in other common accountability models.

An analysis of the responses from the urban school sample data identified the variables that received the highest mean scores from teachers and principals. The responses that received the highest mean average were: utilizing the student SES, awarding bonus points for high growth, and utilizing a growth model to observe student performance. The same variables received the highest mean scores, when comparing the urban school sample data to the whole sample data. The average response data from the urban location type can be found in Table 13.

At the district level, an analysis of the urban district sample identified the variables that received the highest mean score from superintendents. The responses with the highest mean average were: taking the SES into account ($M = 4.33$, $SD = 2.08$), the use of a growth model ($M = 3.67$, $SD = 2.52$), and rewarding districts that show high growth ($M = 3.67$, $SD = 2.31$).

Compared to both the urban school sample data, as well as the total sample data, the superintendent responses were similar.

Table 13

Variable Belief Statements by School Locality–Urban

Variable	*M*	*SD*
Use of ISTEP+ or ECA	2.93	1.36
Growth of bottom 25%	3.71	1.43
Separation of top 75%	3.47	1.27
Penalization for low growth	2.33	1.25
Awarding of bonus for high growth	4.28	1.38
Specific percentage every year	2.44	1.31
Individual student growth over time	4.07	1.32
Inclusion of teacher effectiveness	3.61	1.34
Inclusion of SES status	4.80	1.42
Inclusion of student race	2.63	1.43
Inclusion of student gender	2.28	1.32

Further analysis of the participant responses identified the variables that received the lowest mean scores. Data showed that including student gender into an accountability model was not favored. Utilizing student gender received the lowest mean score among each sample population. Similarly, an accountability system that penalizes a school/district due to low growth received low mean scores from respondents in both the urban school sample and the whole sample. Finally, teachers and principals in the urban school sample expressed that a

system similar to AYP should not be utilized. At the district level, superintendents were in agreement that districts should not be penalized for low growth (M = 1.67, SD = .577). They also denoted that an accountability model similar to AYP should not be utilized (M = 2.00, SD = 1.00). The main difference in average response data was that superintendents did not favor looking at the growth of the bottom 25% (M = 3.00, SD = 1.732).

Because the urban school and district sample responses had similar variables that received the highest and lowest mean scores, continued analysis of the participant data identified variables with the largest mean score change from the whole sample. The largest mean difference occurred with the urban school sample response utilizing ISTEP+ or ECA scores to measure school accountability, evidenced by a decrease of .19 in the mean average over the whole sample response for the same variable. The urban school sample increased .12 in the mean average over the whole sample for awarding bonus points when high growth occurs. A decrease occurred with the urban district sample response for observing growth of the bottom 25%. The urban district sample mean score decreased .68 when compared to the whole sample mean score.

Demographic Statistics by School/District Locality–Suburban

A total of 187 individuals identified their school/districts as suburban. The current position of the respondents included 117 (62.6%) teachers, 56 (29.9%) principals, and 14 (7.5%) superintendents. The response rate for the suburban sample was consistent with the whole sample distribution.

Participants revealed the SES of their schools by indicating the free/reduced lunch rates. For teachers and principals, 77 participants (41.2%) noted 34%–67% free/reduced lunch rate, 61 (32.6%) noted between 0%–33% free/reduced lunch rate, and 34 (18.2%) signified free/reduced

lunch rates between 68%–100%. At the district level, seven (50.0%) noted 34%–67% free/reduced lunch rates, six (42.9%) identified less than 33% free/reduced lunch rates, and one (7.1%) signified a free/reduced lunch rate as being between 68%–100%.

For teachers and principals, 106 (56.7%) participants identified their school accountability grade as an A grade, 33 (17.6%) identified B grade, 16 (8.6%) identified C grade, and 15 (8.0%) indicated a D/F grade. At the district level, 11 (78.6%) participants identified their district accountability grade as A grade, two (14.3%) identified B grade, and one (7.1%) indicated a C district grade.

Descriptive Statistics by School/District Locality–Suburban

Participants were asked questions related to their belief in a specific accountability statement. These statements included distinctive variables that are utilized in the current Indiana accountability system or variables used in other common accountability models.

An analysis of the responses from the suburban school sample data identified the variables that received the highest mean scores from teachers and principals. The responses with the highest mean average were: utilizing the student SES, utilizing a growth model to observe student performance, and awarding bonus points for high growth. The same variables received the highest mean average scores, when comparing the suburban school sample data to the whole sample data. The average response data from the suburban location type can be found in Table 14.

At the district level, an analysis of the suburban district sample data identified the variables that received the highest mean score from superintendents. The responses with the highest mean average were: the use of a growth model ($M = 4.86$, $SD = .770$), rewarding districts that show high growth ($M = 4.36$, $SD = 1.50$), and embedding teacher effectiveness into

an accountability model ($M = 4.36$, $SD = 1.08$). Compared to both the suburban school sample data, as well as the total sample data, the superintendent responses differed with embedding teacher effectiveness into an accountability model.

Table 14

Variable Belief Statements by School Locality–Suburban

Variable	*M*	*SD*
Use of ISTEP+ or ECA	3.18	1.40
Growth of bottom 25%	3.72	1.41
Separation of top 75%	3.43	1.36
Penalization for low growth	2.50	1.30
Awarding of bonus for high growth	4.10	1.42
Specific percentage every year	2.52	1.35
Individual student growth over time	4.13	1.26
Inclusion of teacher effectiveness	3.80	1.29
Inclusion of SES status	4.59	1.34
Inclusion of student race	2.60	1.38
Inclusion of student gender	2.09	1.17

Further analysis of the participant responses identified the variables that received the lowest mean scores. Data showed that including student gender into an accountability model was not favored. Utilizing student gender received the lowest mean score among each sample population. Similarly, an accountability system that penalizes a school/district due to low growth received low mean scores from respondents in both the suburban school sample and the

whole sample. Finally, teachers and principals in the suburban school sample expressed that a system similar to AYP should not be utilized. At the district level, superintendents were in agreement that gender (M = 2.29, SD = 1.44) and race (M = 2.50, SD = 1.35) should not be a variable considered when measuring school accountability. Superintendents also did not favor utilizing a system that penalized districts for low growth (M = 2.86, SD = 1.66). Differences did not exist among the school, district, or total sample.

Because the suburban school and district sample responses had similar variables that received the highest and lowest mean scores, continued analysis of the participant data identified variables with the largest mean score change from the whole sample. The largest mean difference occurred with the suburban school sample response for penalizing schools for low growth, evidenced by a .16 increase in the mean average over the whole sample response for the same variable. The suburban school sample score decreased .14 in the mean average over the whole sample for the inclusion of SES into an accountability model. An increase occurred with the suburban district sample response for utilizing a growth model. The suburban district sample mean score increased .80 when compared to the whole sample mean score.

Demographic Statistics by School/District Locality–Rural

A total of 311 individuals identified their school/districts as rural. The current position of the respondents included 176 (56.6%) teachers, 95 (30.5%) principals, and 40 (12.9%) superintendents. The teacher response rate for rural schools was approximately 7% lower than the whole sample distribution. In addition, the response rate for superintendents was almost 4% higher than the whole sample distribution.

Participants revealed the SES of their schools by indicating the free/reduced lunch rates. For teachers and principals, 194 (62.4%) noted a 34%–67% free/reduced lunch rate, 50 (16.1%)

noted between 0%–33% free/reduced lunch rate, and 26 (8.4%) signified free/reduced lunch rate between 68%–100%. At the district level, 29 (70.7%) noted a 34%–67% free/reduced lunch rate, 11 (26.8%) identified less than 33%, and one (2.4%) signified a free/reduced lunch rate between 68%–100%.

For teachers and principals, 135 (43.4%) participants identified their school accountability grade as A grade, 71 (22.8%) identified B grade, 43 (13.6%) identified C grade, and 18 (5.8%) indicated a D/F grade. At the district level, 18 (43.9%) participants identified their district accountability grade as A grade, 12 (31.7%) identified B grade, nine (22.0%) identified C grade, and one (2.4%) indicated a D or F district grade.

Descriptive Statistics by School/District Locality–Rural

Participants were asked questions related to their beliefs in a specific accountability statement. These statements included distinctive variables that are utilized in the current Indiana accountability system or variables used in other common accountability models.

An analysis of the responses from the rural school sample data identified the variables that received the highest mean scores from teachers and principals. The responses with the highest mean average were: utilizing the student SES, awarding bonus points for high growth, and utilizing a growth model to observe student performance. The same variables received the highest mean average scores, when comparing the rural school sample data to the whole sample data. The average response data from the rural school location type can be found in Table 15.

At the district level, an analysis of the rural district sample data identified the variables that received the highest mean score from superintendents. The responses with the highest mean average were: taking the socioeconomic make-up of the district into account ($M = 4.46$, $SD = 1.36$), looking at the growth of the bottom 25% ($M = 4.15$, $SD = 1.24$), and the use of a growth

model (M = 4.12, SD = 1.17). Compared to both the suburban school sample data, as well as the total sample data, the superintendent responses differed in the use of the bottom 25% growth.

Table 15

Variable Belief Statements by School Locality–Rural

Variable	*M*	*SD*
Use of ISTEP+ or ECA	3.24	1.39
Growth of bottom 25%	3.62	1.25
Separation of top 75%	3.46	1.27
Penalization for low growth	2.24	1.14
Awarding of bonus for high growth	4.14	1.37
Specific percentage every year	2.56	1.30
Individual student growth over time	4.06	1.24
Inclusion of teacher effectiveness	3.67	1.29
Inclusion of SES status	4.74	1.26
Inclusion of student race	2.60	1.43
Inclusion of student gender	2.12	1.19

Further analysis of the participant responses identified the variables that received the lowest mean scores. Data showed that including student gender into an accountability model was not favored. Utilizing student gender received the lowest mean score among each sample population. Similarly, an accountability system that penalizes a school/district due to low growth received low mean scores from respondents in both the rural school sample and the whole sample. Finally, teachers and principals in the rural school sample expressed that a

system similar to AYP should not be utilized. At the district level, superintendents were in agreement that gender (M = 1.85, SD = 1.04) or race (M = 2.27, SD = 1.40) should not be a variable considered when measuring school accountability. Superintendents also did not favor utilizing a system that penalized districts for low growth (M = 2.27, SD = 1.34). No differences existed among the school, district, or total sample.

Because the rural school and district sample responses had similar variables that received the highest and lowest mean scores, continued analysis of the participant data identified variables with the largest mean score change from the whole sample. The largest mean difference occurred with the rural school sample response for utilizing ISTEP+ or ECA scores to measure school accountability, evidenced by an increase of .12 in the mean average over the whole sample response for the same variable. The rural school sample score decreased .10 in the mean average over the whole sample for penalizing schools that showed low growth. An increase occurred with the rural district sample response for observing the growth of the bottom 25%. The rural district sample mean score increased .47 when compared to the whole sample mean score.

Inferential Test Results

Research Questions 2–5 utilized a one-way analysis of variance (ANOVA) to identify if any significant differences existed between the mean scores of three or more groups. This measurement was chosen since each research question had one independent variable with at least three levels and the use of the single dependent variable. Research Question 6 utilized multiple regression analysis. This measurement determined if the predictor variables explained a significant amount of variance within the criterion variable.

The independent variables of this study consisted of educational position type, the IDOE A–F Accountability designation for the 2012–2013 school year, the school or district SES level, and the location of the school or district. The educational position type independent variable had three levels: teacher, principal, or superintendent. The independent variable of location had three levels: suburban, urban, or rural. The independent variable based on SES level had three levels: 0%–33%, 34%–67%, or 68%–100%. The IDOE A–F Accountability designation for the 2012–2013 school year independent variable had four levels: A, B, C, and D/F. The dependent variable was the perception survey composite score as it related to educational position type, location, SES level, and letter grade received from the IDOE during the 2012–13 school year.

For the final research question, the composite score served as the criterion variable. The utilization of five different predictor variables allowed for a multiple regression test to be run. These predictor variables were specific SES level, school enrollment size, building specific accountability grade, location type, and the number of years the principal has been in his or her current position.

Null Hypothesis 1

A one-way ANOVA was used to identify if any significant differences existed between the accountability composite scores of teachers, principals, and superintendents. To ensure the validity of the results, the assumptions for the one-way ANOVA were tested. The accountability composite scores were examined within a box plot to determine if potential outliers existed. The box plot suggested that there were no points of data that were more than 1.5 standard deviations from the box's edge, providing evidence that outliers were not within the model. A Sharpiro-Wilk's test was conducted to check for normality. This assumption was met due to the significance value in the Shapiro-Wilk's test being greater than .05. Finally, homogeneity of

variance was tested utilizing Levene's test of equality of variance. This assumption was met with the significance value, $F = 1.764$, $p = .172$, being greater than .05.

The accountability composite scores for teachers ($M = 18.25$, $SD = 5.02$), principals ($M = 21.21$, $SD = 4.94$), and superintendents ($M = 20.80$, $SD = 5.97$) were examined using a one-way ANOVA to determine if a significant difference existed. The one-way ANOVA found that there were significant differences on the accountability composite score based on the position type, $F(2, 736) = 1.76$, $p < .001$. Due to the significant differences being present, the null hypothesis was rejected. Post hoc testing was required to identify where the significant differences were located.

To determine where the significant difference lies, a Tukey HSD was conducted. The Tukey HSD test detected significant difference. The teacher composite score was significantly lower than both the principal composite score ($p < .001$) and the superintendent composite score ($p = .001$). There was no evidence of significant differences among the principal composite scores and the superintendent composite scores.

Null Hypothesis 2

A one-way ANOVA was used to identify if any significant differences existed in the accountability composite score based on the school's Indiana A–F accountability grade. It was important to note that this test only examined the responses from the teacher and principal participants. The teacher and principal participants could potentially have a different viewpoint due to the fact that the demographic data was for their specific school only. The superintendent's district viewpoint could potentially be affected by different data sets from multiple buildings. Additional testing could be required.

Assumptions were tested to ensure the validity of the results. To determine whether there was an outlier within the dependent variable scores, boxplots were examined. The box plot suggested that there were no points of data that were more than 1.5 standard deviations from the box's edge, providing evidence that outliers were non-existent. A Sharpiro-Wilk's test was conducted to check the assumption of normality. This assumption was met due to the significance value in the Shapiro-Wilk's test being greater than .05. Finally, homogeneity of variance was tested utilizing Levene's test of equality of variance. This assumption was met with the significance value, $F = 1.504$, $p = .212$, being greater than .05.

The accountability composite scores for schools identified as A schools ($M = 19.38$, $SD = 5.43$), B schools ($M = 18.93$, $SD = 5.20$), C schools ($M = 19.27$, $SD = 5.14$), and D/F schools ($M = 19.14$, SD = 4.57) were examined using a one-way ANOVA to determine if a significant difference existed. With the one-way ANOVA, there was no significant difference on the accountability composite score based on the Indiana A–F accountability grade, $F(3, 658) = .258$, $p = .856$. Without any significant differences resulting from the inferential test, the null hypothesis was retained. No further post hoc tests were required due to the lack of significant difference.

Null Hypothesis 3

A one-way ANOVA was used to identify if any significant differences existed in the accountability composite score based on the school's SES. It is important to note that this test only examined the responses from the teacher and principal participants. The teacher and principal participants could potentially have a different viewpoint due to the fact that the demographic data was for their specific school only. The superintendent's district viewpoint

could potentially be affected by different data sets from multiple buildings. Additional testing could be required and is discussed in Chapter 5.

Assumptions were tested to ensure the validity of the results. To determine whether there was an outlier within the dependent variable scores, boxplots were examined. The box plot suggested that there were no points of data that were more than 1.5 standard deviations from the box's edge, providing evidence that outliers were non-existent. In an attempt to further test the assumptions for the ANOVA, a Sharpiro-Wilk's test was conducted to check for normality. This assumption was met due to the significance value in the Shapiro-Wilk's test being greater than .05. Finally, homogeneity of variance was tested utilizing Levene's test of equality of variance. This assumption was met with the significance value, $F = .018$, $p = .982$, being greater than .05.

The accountability composite scores for SES levels 0%–33% ($M = 19.74$, $SD = 5.26$), 34%–67% ($M = 19.05$, $SD = 5.18$), and 68%–100% ($M = 19.02$, $SD = 5.21$) were examined using a one-way ANOVA to determine if a significant difference existed. With the one-way ANOVA, there was no significant difference on the accountability composite score based on SES levels, $F(2, 669) = .862$, $p = .423$. Without any significant differences resulting from the inferential test, the null hypothesis was retained. No further post hoc tests were required due to the lack of significant difference.

Null Hypothesis 4

A one-way ANOVA was used to identify if any significant differences existed in the accountability composite score based on the school's location type. It is important to note that this test only examined the responses from the teacher and principal participants. The teacher and principal participants could potentially have a different viewpoint due to the fact that the demographic data was for their specific school only. The superintendent's district viewpoint

could potentially be affected by different data sets from multiple buildings. Additional testing could be required and is discussed in Chapter 5.

Assumptions were tested to ensure the validity of the results. To determine whether there was an outlier within the dependent variable scores, boxplots were examined. The box plots suggested that there were no points of data that were more than 1.5 standard deviations from the box's edge, providing evidence that outliers were non-existent. In an attempt to further test the assumptions for the ANOVA, a Sharpiro-Wilk's test was conducted to check for normality. This assumption was met due to the significance value in the Shapiro-Wilk's test being greater than .05. Finally, homogeneity of variance was tested utilizing Levene's test of equality of variance. This assumption was met with the significance value, $F = .807$, $p = .447$, being greater than .05.

The accountability composite scores for schools identified as urban ($M = 19.15$, $SD = 5.24$), suburban ($M = 19.44$, $SD = 5.57$), and rural ($M = 19.27$, $SD = 5.10$) were examined using a one-way ANOVA to determine if a significant difference existed. With the one-way ANOVA, there was no significant difference on the accountability composite score based on the location type, $F(2, 725) = .159$, $p = .853$. Without any significant differences resulting from the inferential test, the null hypothesis was retained. No further post hoc tests were required due to the lack of significant difference.

Null Hypothesis 5

A stepwise multiple regression model was conducted to determine if the accountability composite score for principals (independent variable) could be predicted by multiple demographic variables. The dependent variables for this test were the 2012–2013 Indiana A–F accountability letter grade, the specific number of students on free or reduced lunch, the numbers of years a principal had been at a specific school, the specific student enrollment, and the school

locality type. The hypothesis examined whether any of these variables explained a significant amount of variance in the criterion variable.

Prior to the analysis, the data were analyzed to identify any assumption violations. The assumption of independence of residuals was tested by looking at the Durbin-Watson statistic. A Durbin-Watson value "close to the value of two demonstrates that independence of residuals was met" (Lomax & Hahs-Vaughn, 2012, p. 396). In this instance, the Durbin-Watson score was close to two, thus, the assumption was met. The residual plot was examined to measure the assumption of linearity. The relationship between the predictor variables and the outcome variable (composite score) compared to the predicted residuals was linear. Thus, the assumption of linearity was not violated. Continued examination of the plot showed that the data points fell within an absolute value of 2, providing further evidence of linearity. No evidence suggested that the assumption of normality was violated "as the errors shown in the residual plot were normally distributed" (Lomax & Hahs-Vaughn, 2012, p. 412). The residuals were aligned with the normal p-plot versus the predicted values. The assumption of multicollinearity was tested to ensure that the predictor variables were not too closely correlated. Heavily correlated predictor variables would result in the "inability to determine which variable was explaining the variance in the criterion variable" (Lomax & Hahs-Vaughn, 2012, p. 385). The tolerance levels for the predictor variables were analyzed and found to be above the 0.2 minimum needed to satisfy this assumption; thus, multicollinearity was not determined to be a problem with this inferential test. A Levene's test of equality of variances was used to test the homogeneity of variance assumption. The assumption looked to ensure that variances within both groups of the dependent variable were equal. This assumption was met with the significance value, $F = 1.78$, $p = .118$, being greater than .05.

An analysis of the multiple correlation coefficient between the observed and predicted values of the criterion variable (composite score) reported a correlation coefficient of .215, "suggesting a weak correlation" (Lomax & Hahs-Vaughn, 2012, p. 377). The coefficient of multiple determination (R^2) was .046, meaning that 4.6% of the variance within the composite scores were explained by the set of predictors. After adjustment was made for the sample size and the number of predictors, the adjusted R^2 was .020 or 2.0% variance explained. This accounted for a .026 amount of shrinkage within the model. Reflecting the "average residual difference in the data points from the regression line", the standard error of the estimate" (4.87) served as the standard deviation for the model" (Lomax & Hahs-Vaughn, 2012, p. 396).

When "adjusted for the number of predictors and for sample size", the variance decreased to 2.0% (Lomax & Hahs-Vaughn, 2012, p. 396). To determine the significance of this finding, an ANOVA was conducted. The ANOVA results can be found in Table 16.

Table 16

ANOVA for the Regression Model for Demographic Factors Predicting the Composite Score of Principals

Model	Sum of Squares	*df*	Mean Square	*F* Value	Sig.
Regression	211.37	5	42.74	1.78	.118

In order to assess the significance of the coefficient of multiple determination (R^2) in the model, an ANOVA test was conducted. The test determined that demographic factors do not serve as significant predictors of the composite score for principals at the 0.05 priori alpha level. The ANOVA was not significant; thus, the original null hypothesis was retained.

Summary

In this chapter, quantitative analysis was utilized to examine the six research questions in this study. The first four null hypotheses were tested by conducting one-way ANOVAs through SPSS. The ANOVA test was completed to determine if there were any significant differences in the accountability composite scores involving the independent variables of position type, the 2012-2013 Indiana A–F accountability grade, SES of the school, and the school locality. Only one of the null hypotheses were rejected. A significant difference existed in the composite scores when separated by position type. The composite score of the teacher sample population was significantly lower than the composite scores of the principals and superintendents.

Three of the null hypotheses were retained meaning there was no significant difference in the accountability composite scores based on the 2012–2013 Indiana A–F accountability grade, the SES of the school, or the location description of the school. As stated in this chapter, the tests for Null Hypotheses 2 through 4 only utilized the teacher and principal composite scores as it was specific to a single school. The superintendent data were not utilized during this process due to the possibility of multiple data sets within a school district/community.

A multiple regression test was used to analyze Null Hypothesis 5. This test attempted to predict the principal composite score utilizing specific predictor variables. The predictor variables utilized were the 2012–2013 Indiana A–F accountability grade, the specific number of students on free or reduced lunch, the specific school enrollment size, school location, and the number of years the principal had been at his or her specific school. Analysis of Null Hypothesis 5 showed that none of the five predictor variables explained a significant amount of variance in the composite score.

The Perceptions of Variables Utilized to Effectively Measure School Accountability Survey contained two open-ended survey questions relating to accountability practices. The practice of measuring school accountability garners mixed perceptions. Teachers, principals, and superintendents supplied alternative assessment options and reviews through the open-ended survey questions. The implications discussed in Chapter 5 were developed through an analysis of the open-ended survey responses.

CHAPTER 5

SUMMARY OF FINDINGS, RESULTS, IMPLICATIONS AND RECOMMENDATIONS

This chapter is organized into four sections that include the summary of findings, results, implications, and recommendations for future research. The summary section addresses the purpose of the study and an overview of what this study discovered. The results section summarizes the specific results of each research question. The implications section summarizes what impact this study may have on current and future accountability practices as it relates to education. Finally, the recommendation for future research summarizes the possibility of future research in the area of school accountability.

Summary of Findings

This study examined a portion of the NCLB accountability system, along with alternative accountability variables, in an attempt to identify specific variables members of the education profession believe should be utilized. Linn (2008) argued that a state's ability to identify a successful school or a failing school may be inconsistent due to the methodology that is utilized by the state to measure AYP rather than the actual effectiveness of schools. This study focused on the variables that are utilized to measure school accountability and their significance as indicators of school performance. Additionally, this study sought to identify specific demographic factors that might predict a principal's choice in variables used to measure school accountability. The study's intent was to provide educators, educational leaders, and legislators

with information that could be utilized when creating or adapting current accountability models. The study was designed with the intent of allowing professional educators an opportunity to voice their beliefs in specific variables used to measure school accountability. Specifically, it was the intention of this research to help provide data to support future decision-making practices of those educators and/or legislators concerned with school accountability.

This quantitative, survey-based research was used to determine teachers', principals' and superintendents' perceptions of the variables needed to effectively measure school accountability. Also investigated were the responses from these three groups in relation to four different criteria (a) perception based on position type, (b) perception based on the school/district's Indiana accountability grade, (c) perception based on the socioeconomic make-up of the school/district, and (d) perception based on the school/district's location type. The quantitative design of this study allowed for statistical models to be run answering each research question. A total of 746 individuals participated in this survey. The total sample for this study consisted of 471 (63.1%) respondents who identified as teachers, 209 (28.0%) principals, 59 superintendents, and seven (.9%) respondents who did not identify their current position.

Results

Research Question 1

Research Question 1 asked, "What are the current perceptions of Indiana educators on the variables in various accountability systems?" The top three whole sample responses were the inclusion of a student's SES into the accountability model, the ability to gain bonus points for showing positive growth, and, observing individual student growth from one year to the next. Further analysis was conducted to identify the variables that contained the least amount of belief. The whole sample response showed that including student gender into an accountability model

was not favored. In terms of how student performance data was utilized, an accountability system which penalizes a school for low growth garnered little trust. Finally, the whole sample did not favor having a specific yearly target, similar to AYP, that students were expected to meet.

Research Question 2

Research Question 2 asked, "Is there a significant difference in the accountability composite score based on the type of position of the respondents?" The accountability composite scores for teachers (M = 18.25, SD = 5.02), principals (M = 21.21, SD = 4.94), and superintendents (M = 20.80, SD = 5.97) were examined using a one-way ANOVA to determine if a significant difference existed between position types. The one-way ANOVA found significant differences on the accountability composite score based on the position type, $F(2, 736) = 1.76$, $p < .001$. Because the significant differences were present, the null hypothesis was rejected. The Tukey HSD post hoc test detected significant difference among the composite scores of the teacher sample. The teacher composite score was significantly lower than both the principal composite score and the superintendent composite score. The teacher composite score was significantly lower than both the principal composite score ($p < .001$) and the superintendent composite score ($p = .001$). Table 17 shows a comparison of mean scores between teachers, principals, and superintendents.

Table 17

Comparison of Mean Composite Scores Between Position Type with Significant Results

Position	Mean Score	*SD*	Mean Score Difference (From Teacher)
Teacher	18.25	5.02	
Principal	21.21	4.94	-2.96*
Superintendent	20.80	5.97	-2.53*

Note. *$p \leq .05$

Research Question 3

Research Question 3 asked, "Is there a significant difference in the accountability composite score based on the school or district's accountability grade?" A one-way ANOVA was used to identify if any significant differences existed in the accountability composite score based on the school's Indiana A–F accountability grade? It is important to note this test only examined the responses from the teacher and principal participants. The teacher and principal participants could potentially have a different viewpoint due to the fact that the demographic data was for their specific school only. The superintendent's district viewpoint could potentially be affected by different data sets from multiple buildings. Additional testing could be required and is discussed in this chapter.

The accountability composite scores for school identified as A schools ($M = 19.38$, $SD = 5.43$), B schools ($M = 18.93$, $SD = 5.20$), C schools ($M = 19.27$, $SD = 5.14$), and D/F schools ($M = 19.14$, $SD = 4.57$) were examined using a one-way ANOVA to determine if a significant difference existed. With the one-way ANOVA, no significant difference existed on the accountability composite score based on the Indiana A–F accountability grade, $F(3, 658) = .258$,

$p = .856$. Without any significant differences resulting from the inferential test, the null hypothesis was retained.

Research Question 4

Research Question 4 asked, "Is there a significant difference in the accountability composite score based on the school or district's socioeconomic level?" A one-way ANOVA was used to identify if any significant differences existed in the accountability composite score based on the school's SES? It is important to note this test only examined the responses from the teacher and principal participants. The teacher and principal participants could potentially have a different viewpoint due to the fact that the demographic data is for their specific school only. The superintendent's district viewpoint could potentially be affected by different data sets from multiple buildings. Additional testing could be required and is discussed later in Chapter 5.

The accountability composite scores for socioeconomic levels 0 – 33% ($M = 19.74$, $SD = 5.26$), 34 – 67% ($M = 19.05$, $SD = 5.18$), and 68 – 100% ($M = 19.02$, $SD = 5.21$) were examined using a one-way ANOVA to determine if a significant difference existed. With the one-way ANOVA, no significant difference arose on the accountability composite score based on socioeconomic levels, $F(2, 669) = .862$, $p = .423$. Without any significant differences resulting from the inferential test, the null hypothesis was retained.

Research Question 5

Research Question 5 asked, "Is there a significant difference in the accountability composite score based on the school or district's location type?" A one-way ANOVA was used to identify if any significant differences existed in the accountability composite score based on the school's location type? It is important to note this test only examined the responses from the teacher and principal participants. The teacher and principal participants could potentially have a

different viewpoint due to the fact that the demographic data was for their specific school only. The superintendent's district viewpoint could potentially be affected by different data sets from multiple buildings. Additional testing could be required and is discussed later in this chapter.

The accountability composite scores for schools identified as urban (M = 19.15, SD = 5.24), suburban (M = 19.44, SD = 5.57), and rural (M = 19.27, SD = 5.10) were examined using a one-way ANOVA to determine if a significant difference existed. With the one-way ANOVA, no significant difference was determined on the accountability composite score based on the location type, $F(2, 725) = .159$, $p = .853$. Without any significant differences resulting from the inferential test, the null hypothesis was retained.

Research Question 6

Research Question 6 asks, "Does school socioeconomic status level, school enrollment size, years in current position as a principal, location, and school accountability grade serve as predictors of the composite score? A stepwise multiple regression model was conducted to determine if the accountability composite score for principals (independent variable) could be predicted by multiple demographic variables. The dependent variables for this test were: the 2012 – 2013 Indiana A – F accountability letter grade, the specific number of students on free or reduced lunch, the numbers of years a principal has been at a specific school, the specific student enrollment, and the school locality type. The hypothesis examined whether any of these variables explain a significant amount of variance in the criterion variable.

An analysis of the multiple correlation coefficient between the observed and predicted values of the criterion variable (composite score) reported a correlation coefficient of .215, "suggesting a weak correlation" (Lomax & Hahs-Vaughn, 2012, p. 377). The coefficient of multiple determination (R^2) was .046, meaning that 4.6% of the variance within the composite

scores are explained by the set of predictors. After adjustment was made for the sample size and the number of predictors, the adjusted R^2 was .020 or 2.0% variance explained. This accounted for a .026 amount of shrinkage within the model. Reflecting the "average residual difference in the data points from the regression line", the standard error of the estimate (4.87) served as the standard deviation for the model (Lomax & Hahs-Vaughn, 2012, p. 396).

In order to assess the significance of the coefficient of multiple determination (R^2) in the model, an ANOVA test was conducted. The test determined that demographic factors did not serve as significant predictors of the composite score for principals at the 0.05 priori alpha level. The null hypothesis was retained.

Implications

With the increased utilization of student performance as a means of holding educational institutions accountable, as witnessed within NCLB, states should have a reliable accountability system that can accurately and fairly measure schools. The impact of NCLB legislation continues to grow, and the continued reauthorization of specific provisions within NCLB proves that it will continue to be a part of accountability within education for the future.

Implication 1: Current Indiana Accountability Practice

Relative to Research Question 1, specific questions within the survey assessed current accountability practices in Indiana, and specific questions assessed common accountability practices elsewhere. In regard to current practices in the state of Indiana, the perceptions of specific variables were mixed. For example, 53% of teachers, principals, and superintendents believed that achievement results from ISTEP+ or ECA should not be utilized to measure school accountability. Per the NCLB (2002) waivers, the current practice in the state of Indiana is that the results from standardized testing have a major impact on a school's accountability grade.

This study also showed that 62% of teachers, principals, and superintendents felt that it was important to look at the growth of the bottom 25% of students. Again, this is a current practice in the state of Indiana. However, close to 81% of these same participants shared that they disagreed with penalizing schools or districts for having a large population of students showing low growth. Conversely, 57% of the respondents felt that schools or districts should receive bonus points for showing high growth.

This study would suggest that the current practice in Indiana of observing student growth was favored, as nearly 75% of teachers, principals, and superintendents showed support for this model. This study would imply that the current Indiana practice of utilizing a growth model that would potentially award schools bonus points for showing high growth, yet not penalize schools for low growth would be favored. However, choosing what data should be observed over a period of time would still be in question.

Implication 2: Accountability Practices Outside Indiana

The study also surveyed participants about accountability practices utilized outside of the state of Indiana. The federal government uses adequate yearly progress to measure school accountability. Seventy-two percent of the educators in Indiana shared that they disagreed with the use of an accountability model similar to AYP.

Different examples of accountability models also utilize different variables to measure accountability. Teacher effectiveness is an example of a variable that has gained in popularity. The ability to pinpoint the effects of a specific teacher or school has prompted policymakers to look at a VAM as an opportunity to connect student achievement to accountability. Some districts, such as New York City, have even begun using a VAM to bridge student performance to educator pay. Advocates for VAMs argue that this could shed new light on individual teacher

accountability; however, differences in opinion and research exist (Goldschmidt et al., 2005). This survey showed that 66% of the educators in the state of Indiana believed that teacher effectiveness should be included in the design of an accountability model.

Another variable that is not currently utilized in the state of Indiana is SES. Nearly 86% of the teachers, principals, and superintendents shared that they would be supportive of an accountability model that considered the socioeconomic status of students. However, nearly 74% of the participants felt that students' race and 85% of the participants felt that students' gender should not be considered when measuring accountability. These final two variables, race and gender, have been utilized in VAMs around the country.

Implication 3: Potential Changes to the Indiana Accountability Model

Based specifically on the survey results, this study would imply that the current accountability model in the state of Indiana could be altered. This information makes clear the implications for state legislators. As identified above, a growth model is the preferred accountability model. However, Indiana educators voiced their support of a more detailed growth model that utilized multiple variables to measure accountability, similar to a VAM. It would be perceived that a value-added growth model that identified struggling learners, as well as proficient learners would be preferred. When measuring the growth of students, the student background, prior achievement, and other data would be observed to identify the specific effects of a school, teacher, or program on student academic progress (Goldschmidt et al., 2005; Green, 2010). The potential VAM would allow for a school to be rewarded for showing high levels of growth, but would not be penalized if students showed low growth.

Implication 4: Effect of Demographic Factors

Relative to Research Questions 2, 3, 4, and 5, this study showed that the perception of

accountability factors was not altered by specific demographic data. It would have been easy to assume that demographic information such as location type, SES, Indiana A–F accountability grade, and position type could have an influence on the belief about specific variables used to measure school accountability. For example, it could be assumed that a school or district with a small number of students who received free or reduced lunch would not feel as strong about including this variable into the accountability model as a school or district that had a high proportion of students who received this service. The results of this survey showed that, regardless of position type, previous Indiana A – F accountability letter grade, socioeconomic make-up and location, consistent feelings about what type of data and how to effectively utilize this data to measure school accountability existed.

Implication 5: Need for Continued Discussion

This study also showed different perceptions in regard to school accountability. For the past 13 years, NCLB has been a critical topic in the educational profession. Under NCLB, schools are held accountable through a measurement of student achievement levels in both the school and school district. The philosophical focus on every student meeting ambitious academic goals to close the achievement gap under NCLB has been commended (Borowski & Sneed, 2006; Chubb, 2005; Guilfoyle, 2006; Haycock, 2006; Hess, 2006; Hess & Petrilli, 2006; Lewis, 2006; Staiger et al., 2002). According to Guskey (2007), educational leaders should embrace accountability that focuses on student achievement to improve the quality of education in school, as a positive aspect of NCLB.

Proponents of NCLB argue that this relentless insistence of 100% proficiency gives educators, parents, and students a common goal. It is also argued that educational institutions that are failing students should be held accountable. In contrast, opponents of NCLB claim that

this law actually leaves students even farther behind. Overall, as supported by the results of Research Question 2, this study showed that teachers have a statistically significantly lower perception of school accountability when compared to principals and superintendents. The results of this study indicate a need for the educational community to do a better job of building the capacity of knowledge of its professionals in regard to accountability. Improving students and schools should be a central focal point of every educator. By opening up conversations and coming to an agreement on how to effectively measure this progress is important. In a 2012 study, Vernaza found

> Teachers have voices that must be recognized at the policy level, and possess valuable data insights that are worthy of further discussion. Data emphasized the need for dialogue between policymakers and educators, and suggests that educators should have a voice in improving accountability systems so that they can assume greater ownership of high-stakes assessment policies. (p. 12)

As educational leaders and legislators continue to look for ways to assess school accountability, this research shows that change in the current model should take place. By utilizing the results of this research, I would recommend that the state of Indiana continue to pursue an accountability that measures student growth over a period of time. The growth model allows for the results of individual students to be measured, instead of looking at students as part of a group. It is also recommended that the growth model become even more specific than it is in its current state. Having the ability to remove external factors, such as SES, would allow evaluators to pinpoint the positive or negative effect that a teacher, school or program has on student learning. More importantly, I recommend that state legislators and educational leaders utilize the expertise and opinions of professionals that work alongside students. By allowing

teachers, principals, and superintendents to have a voice in the accountability discussion, a system can be created that will not only garner the support of all, but will also truly monitor the progress of education in Indiana.

Implication 6: Indiana's Proposed Accountability Model

In 2013, Indiana legislators requested a review of the current accountability measurement system. Representatives from around the state proposed a new accountability model aimed to accurately and fairly measure schools. The 2015 State Board of Education adopted a plan altering the way school letter grades were assigned. Indiana's proposed student-centered accountability system will calculate school grades based on three dimensions: performance, growth, and multiple measures. The growth indicator will observe individual student data from year-to-year to calculate scores in Grades 3–10. The performance indicator will utilize specific student passing and participation rates in both English/language arts and mathematics. For high schools, the graduation rates, along with the college and career readiness scores, make up the final multiple-measure indicators. Schools will receive a final total score average based on a 100% to determine the school letter grade (Ritz & Yager, 2013).

The major change from the 2012 model to the proposed 2015 model was the measurement of growth. The newly proposed accountability model will measure individual student growth. This is in contrast to the current model that measures the growth relative to the cohorts of students scoring the same score in the previous year. The proposal is an attempt to compare the academic performance and growth of the individual student in each school with the prior academic performance and growth of the individual students in the school or corporation and not to the performance of other schools or corporations (Ritz & Yager, 2013). The shift to observing individual student growth is a variable supported by teachers, principals, and

superintendents in the research study. Along with individual student growth, schools are no longer awarded bonus points for showing high growth. This specific variable received mixed mean scores among teachers, principals, and superintendents. The bonus points could have been viewed as an incentive to schools to maintain high growth; however, the newly proposed utilizes multiple placement categories as a means to incentivize progress. Student growth is relative to how their proficiency scores will fall within eight categories from year to year. The yearly student movement within categories, or lack thereof, will determine the points awarded to each school. This correlates with the negative perception of penalizing schools that do not show high growth by allowing schools to receive specific points for any and all individual student growth. Achieving a years worth of growth allows a student to attain more points for a school than a student exhibiting modest growth.

Data from my research study suggests the newly proposed student-centered accountability model is a step in the right direction. The variables receiving the highest mean scores in this research study were considering in the newly proposed accountability model, with the exception of socio-economic status. As discussed in this research study, poverty continues to be a variable that has an impact on student learning. It would seem that legislators do not accept poverty as a reason for low student performance. The student-centered accountability model attempts to accommodate this variable by utilizing different proficiency categories. However, until a model can truly separate the effect a school has on student performance from external variables, such as poverty, measuring school accountability will be challenging.

Areas for Further Research

This study focused on the perceptions of accountability by Indiana teachers, principals, and superintendents. During the final stages of this study, additional research studies were

identified that would lead to further exploration of accountability.

I recommend that researchers conduct a quantitative study in which the superintendent responses are separated from the teacher and principal responses to measure the effect of demographic factors on district composite scores. This study would allow the perceptions of superintendents to be measured based at the district level. The original study attempted to measure these perceptions; however, the results could not be compared against individual schools. The fact that a district could potentially consist of multiple schools meeting different demographic factors, a study looking only at superintendents could give us better insights into their perceptions of variables used to measure school accountability.

The high-stakes testing and accountability that accompanied NCLB (2002) legislation has generated interest as well as data around the topic of accountability effects in the nation's public school system. My recommendation is that a complete meta-analysis of accountability models be conducted to see if state accountability efforts have succeeded in raising student achievement.

It would also be insightful to conduct a mixed design study in which legislators on both the Indiana House and Senate Education Committees share their opinion through qualitative and quantitative methods in order to get a more thorough understanding of accountability variables and models. Recent history has shown that legislators are becoming more involved in writing educational policy. Currently, 25% of the bills that have been introduced in the Indiana House and Senate are related to schools (Indiana General Assembly, 2015). Conducting this research would give insight to the perceptions and beliefs of the individuals writing this policy.

Conducting a study utilizing specific student achievement data that is run through various accountability models to observe any discrepancies in accountability outcomes would be beneficial. This study would allow the public to see how different models and the utilization of

different variables potentially impact the accountability designation. This study would highlight the best procedures for determining effective/ineffective schools and identify variables that have consistent outcomes for all Indiana schools.

As changes occur in the current accountability system in Indiana, it would be recommended that this study be replicated in the future to see if the addition of new variables or accountability models can change the perceptions of teachers, principals, superintendents. Current reform in our accountability model continues to fine-tune the process for measuring schools. A replication of this study, during this reform, will monitor the perceptions of educators around the state.

Finally, I would recommend conducting a mixed-design study in which educational professionals share their opinions through quantitative and qualitative methods in order to get a more thorough opinion of accountability measurement. The survey method allowed for mass data to be collected. However, I believe it would be beneficial to conduct interviews with specific professionals in the education system. The opportunity to "tell their story" might allow for a further understanding of their beliefs and perceptions.

Summary

With the increased utilization of student performance as a means of holding educational institutions accountable, as witnessed within NCLB, states should have a reliable accountability system that can accurately and fairly measure schools. The impact of NCLB legislation continues to grow, and the continued reauthorization of specific provisions within NCLB proves that it will continue to be a part of accountability within education for the future. This information could be useful in determining if NCLB reform, and its flexibility, is meeting the goals that were set out by the legislature at both the federal and state levels. This study has

provided the reader with valuable information and insight regarding the perception of accountability and how it is measured. As individual student achievement data continues to be collected on students, the variables and methods in which schools are measured will continue to change. The belief that the need to focus on larger groups of students continues to shift toward the focus on individual student achievement. Regardless of what data is observed, there needs to be an effective way to utilize the data to measure the progress of a school or district. I recommend that legislators and/or school leaders embrace the need to include all stakeholders in this decision.

Legislators will continue to find opposition in a system that is perceived as unfair or "handed down" to educators. The responses in Appendix C show that there is much apprehension in past and present accountability practices. This data does not include all of the teachers, principals, and superintendents that chose not to participate in this study for various reasons. The challenge of legislation will be to identify a viable accountability option that is agreed upon by all stakeholders that can effectively measure the progress of schools for years to come. Identifying schools that are ineffective and finding ways to assist schools in improving is at the heart of accountability. Like it or not, accountability needs to exist to make certain that students are progressing. Educators, as well as legislators, need to work together to find an effective system that fairly hold schools and districts accountable for student learning.

REFERENCES

Arnold, M. L. (2000). *Rural schools: Diverse needs call for flexible policies.* [Policy Brief]. Aurora, CO: Midcontinent Research for Education and Learning. Retrieved from http://files.eric.ed.gov/fulltext/ED441656.pdf

Baker, E. L. (2002, April 1-5). *Visions of test results dance in their heads*. American Educational Research Association Conference, New Orleans, LA. Retrieved from http://www.cse.ucla.edu/products/overheads/aera2002/Vsns02AERAelb.pdf

Bancroft, K., Barker, R. G., & Gump, P. V. (1964). *Big school, small school: High school size and student behavior.* Palo Alto, CA: Stanford University Press.

Bankston, C., & Caldas, S. (1998). Family structure, schoolmates, and racial inequalities in school achievement. *Journal of Marriage & the Family, 60,* 715-724. Retrieved from http://www.jstor.org/stable/353540

Barth, R. S. (1990). *Improving schools from within.* San Francisco, CA: Jossey-Bass.

Beck, F. D., & Shoffstall, G. W. (2005). How do rural schools fare under a high-stakes testing regime? *Journal of Research in Rural Education, 20*(14). Retrieved from http://jrre.psu.edu/articles/20-14.pdf

Bennett, T. (2011, December 7). *State Board of Education monthly meeting* [Video recording]. Retrieved from http://www.doe.in.gov/idoe/sboe/minutes-and-streaming-video

Bergeson, T. (2006). *Race, poverty and academic achievement.* Retrieved from http://www.doh.wa.gove/SBOH/ESS/documents/Race&Poverty.pdf

Borowski, J. W., & Sneed, M. (2006). Will NCLB improve or harm public education? *Harvard Educational Review, 76,* 457-460. Retrieved from http://hepg.org/her-home/issues/harvard-educational-review-volume-76-issue-4/herarticle/_192

Bradley, R. H., Corwyn, R. F., McAdoo, H. P., & Coll, C. G. (2001). The home environments of children in the United States, Part I: Variations by age, ethnicity, and poverty status. *Child Development, 72,* 1844–1867. doi:10.1111/1467-8624.t01-1-00383

Brennan, R. L. (2001). Some problems, pitfalls, and paradoxes in educational measurement. *Educational Measurement: Issues and Practice, 20*(6), 6-18. doi:10.1111/j.1745-3992.2001.tb00071.x

Bryk, S. A., Thum, M. Y., Easton, Q. J., & Luppescu, S. (1998). *Academic productivity of Chicago public elementary schools.* Chicago, IL: Consortium on Chicago School Research.

Carter, S. (1999*). No excuses: Seven principals of low-income schools who set the standard for high achievement.* Washington, DC: The Heritage Foundation.

Choi, K., Goldschmidt, P., & Yamashiro, K. (2006). *Exploring models of school performance: From theory to practice* [CSE Rep: No 673]. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing. Retrieved from http://files.eric.ed.gov/fulltext/ED518678.pdf

Choi, K., Seltzer, M., Herman, J., & Yamashiro, K. (2004). *Children left behind in AYP and non-AYP schools: Using student progress and the distribution of student gains to validate AYP* [CSE Rep: No. 637]. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing. doi:10.1111/j.1745-3992.2007.00098.x

Chubb, J. E. (2005). Saving No Child Left Behind. In J. E. Chubb (Ed.), *Within reach: How America can educate every child* (pp. 1-34). Lanham, MD: Rowman & Littlefield.

Civil Rights Act of 1964, Pub.L. 88-352, 78 Stat. 241. (1964).

Clark, D., Martorell, P., & Rockoff, J. (2009). *School principals and school performance.* Washington, DC: National Center for Analysis of Longitudinal Data in Education Research. Retrieved on January 16, 2012, from http://www.caldercenter.org/upload/Working-Paper-38_FINAL.pdf

Coleman, S. (1966). *Equality of Educational Opportunity (COLEMAN) Study (EEOS).* ICPSR06389-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2007-04-27. doi:10.3886/ICPSR06389.v3

Colorado Department of Education. (2011, September 23). *Colorado's education commissioner joins president in announcing waivers to NCLB law today* [Press release]. Retrieved from http://www.cde.state.co.us/communications/releases/20110923president

Colorado Department of Education. (2012, February 9). *Colorado among first in nation to be granted waivers from NCLB* [Press release]. Retrieved from http://www.cde.state.co.us/communications/Releases/20120209nclbwaivergranted.html

Colorado Department of Education. (2014). *District accountability handbook* (Version 4). Retrieved from http://www.cde.state.co.us/sites/default/files/District%20 Accountability%20Handbook2013v4.pdf

Conant, J. B. (1959). *The American high school today.* New York, NY: McGraw-Hill.

Cox, J., & Cox, K. B. (2008). *Your opinion please! How to build the best questionnaires in the field of education* (2nd ed.). Thousand Oaks, CA: Corwin Press.

Creswell, J. (2003). *Research design: Qualitative, quantitative, and mixed method approaches* (2nd ed.). Thousand Oaks, CA: Sage.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed method approaches.* Thousand Oaks, CA: Sage.

Crouch, R. (2012). *The United States of education: The changing demographics of the United States and their schools.* Alexandria, VA: Center for Public Education. Retrieved August 20, 2014, from http://www.centerforpubliceducation.org/You-May-Also-Be-Interested-In-landing-page-level/Organizing-a-School-YMABI/The-United-States-of-education-The-changing-demographics-of-the-United-States-and-their-schools.html

CTB/McGraw-Hill. (2002). *Accountability and educational progress: A primer on the assessment provisions of the new ESEA.* Retrieved August 13, 2013, from http://www.ctb.com/media/articles/pdfs/general/ctb_esea_primer.pdf

Cuban, L. (2004). Looking through the rearview mirror at school accountability. In K. A. Sirotnik (Ed.), *Holding accountability accountable: What ought to matter in public education* (pp. 18-34). New York NY: Teachers College Press.

Darling-Hammond, L. (1999). *Teacher quality and student achievement: A review of state policy evidence* (Document R-99-1). Retrieved December 26, 2014, from http://www.ctpweb.org

Dimmock, C., & Lee, J. C. (2000). Redesigning school-based curriculum leadership: A cross-cultural perspective. *Journal of Curriculum and Supervision, 15,* 332-358. Retrieved from http://www.ascd.org/publications/jcs/summer2000/Redesigning_School-Based_Curriculum_Leadership@_A_Cross-Cultural_Perspective.aspx

Dokoupil, T. (2010, November 25). New study: School consolidation gets failing grade. *Newsweek.* Retrieved April 21, 2014, from http://www.newsweek.com/new-study-school-consolidation-gets-failing-grade-74145

Donahue, P., Voelkl, K., Campbell, J., & Mazzeo, J. (1999). *NAEP 1998 reading report card for the nation* (NCES Report No. 1999-459). Washington, DC: National Center for Educational Statistics, Office of Educational Research and Improvement

Doran, H. C., & Izumi, L. T. (2004). *Putting education to the test: A value-added model for California.* Chicago, IL: Heartland Institute. Retrieved May 18, 2014, from http://heartland.org/policy-documents/putting-education-test-value-added-model-california

Duran, A. (2005). Factors to consider when evaluating school accountability results. *Journal of Law and Education, 34*(1), 73-100. Retrieved from http://www.ppmrn.net/storage/ppmrn/factors.pdf

Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership, 1(37)*, 15-24. Retrieved from http://larrycuban.files.wordpress.com/2012/10/el_107910_edmonds-2.pdf

Educate America Act. (1993). H.R. 1804 - 103rd Congress: Goals 2000. Retrieved December 15, 2013 from http://www.govtrack.us/congress/bills/103/hr180

Elementary and Secondary Education Act, 20 USC 6651. (1965).

Elliot, S. (2012, February 9). Some give Indiana's new grading system failing marks. *The Indianapolis Star.* Retrieved from www.indystar.com

Elmore, R. F. (2002). Unwarranted intrusion. *Education Next, 2*(1), 30-35. Retrieved from http://educationnext.org/unwarranted-intrusion/

Erpenbach, W., Forte-Fast, E., & Potts, A. (2003). *Statewide educational accountability under NCLB. Accountability systems and reporting state collaborative on assessment and student standards paper*. Washington, DC: Council of Chief State School Officers.

Erpenbach, W. J., & Forte, E. (2005). *Statewide educational accountability under the No Child Left Behind Act: A report on 2005 amendments to state plans.* Washington, DC: Council of Chief State School Officers. Retrieved from http://www.ccsso.org/Documents/2007/Statewide_Educational_Accountability_Systems_2007.pdf

Evers, W. M., & Walberg, H. J. (Eds.) (2002). *School accountability.* Stanford, CA: Stanford University, Hoover Institution Press.

Finn, C. E., Jr. (1989). A nation still at risk. *Commentary,* 17-23. Retrieved from http://www.unz.org/Pub/Commentary-1989may?View=SummaryTree&OpenNodes=Commentary,-198,-1989,-1989may

Fowler, W. J., & Walberg, H. J. (1991). School size, characteristics, and outcomes. *Educational Evaluation and Policy Analysis, 13,* 189-202. Retrieved from http://eric.ed.gov/?id=EJ430692

Fujino, G. (2013). *Evaluating teachers with student growth percentiles and aims proficiency in the state of Arizona* (Doctoral dissertation, Northern Arizona University). Retrieved from http://gradworks.umi.com/35/62/3562126.html

Glidden, H. G. (1999). Breakthrough schools: Characteristics of low-income schools that perform as though they were high-income schools. *ERS Spectrum, 17*(2), 21-26. Retrieved from http://eric.ed.gov/?id=EJ589452

Goals 2000: Reforming Education to Improve Student Achievement - History - April 30, 1998, (n.d.) Retrieved from http://www2.ed.gov/pubs/G2KReforming/g2ch1.html

Goertz, M. E., & Duffy, M. C. (2001). Factors to consider when evaluating school accountability results (CPRE Policy Brief). *Consortium for Policy Research in Education, 1,* 1-7. Retrieved from http://www.ppmrn.net/storage/ppmrn/factors.pdf

Gold, D. L. (1988, April 13). Extensive tests in Indiana make teachers uneasy. *Education Week.* Retrieved from http://www.edweek.org/ew/articles/1988/04/13/29istep.h07.html?r=1571667222

Goldberg, M. K. (2000). An interview with John Goodlad: Leadership for change. *Phi Delta Kappan, 82*(1), 82-85. Retrieved from http://eric.ed.gov/?id=EJ612902

Goldschmidt, P. (2004, June). *The relationship between school quality and the probability of passing standards-based high-stakes performance assessments*. Paper presentation, Symposium Large Scale Assessment Conference, Boston, MA.

Goldschmidt, P., Roschewski, P., Choi, K., Auty, W., Hebbler, S., Blank, R., & Williams, A. (2005). *Policymakers' guide to growth models for school accountability: How do accountability models differ?* A paper commissioned by the Council of Chief State School Officers, Washington, DC. Retrieved from http://www.ccsso.org/Documents/2005/Policymakers_Guide_To_Growth_2005.pdf

Gong, B., Perie, M., & Dunn, J. (2006). *Using student longitudinal growth measures for school accountability under No Child Left Behind: An update to inform design decisions.* Accessed May 19, 2014 from http://www.nciea.org/publications/GrowthModel Update_BGMAPJD07.pdf

Goodlad, J. I. (1984). *A place called school: Prospects for the future.* New York, NY: McGraw- Hill.

Green, J. (2010). *Estimating teacher effects using value-added models* (Doctoral dissertation). Retrieved from http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1005&context=statisticsdiss

Gregory, T., & Smith, G. (1987). *High schools as communities: The small school reconsidered.* Bloomington, IN: Phi Delta Kappa Educational Foundation. Retrieved from http://files.eric.ed.gov/fulltext/ED278518.pdf

Guilfoyle, C. (2006). NCLB: Is there life beyond testing? *Educational Leadership, 64*(3), 8-13. Retrieved from http://www.ascd.org/publications/educational-leadership/nov06/vol64/num03/NCLB@-Is-There-Life-Beyond-Testing%C2%A2.aspx

Guskey, T. (2007). Leadership in the age of accountability. *Educational Horizons, 86*(1), 29-34. Retrieved from http://files.eric.ed.gov/fulltext/EJ776908.pdf

Hager, G. (2006). *School size and student outcomes in Kentucky's public schools.* (Research Report No. 334). Frankfort, KY: Legislative Research Commission. Retrieved from http://www.lrc.ky.gov/lrcpubs.RR334.pdf

Hannaway, J. (1997). *Reports of reform from the field: District and state survey results: Final report.* Washington, DC: U.S. Department of Education, Office of the Under Secretary.

Harkreader, S., & Weathersby, J. (1998). *Staff development and student achievement: Making the connection in Georgia schools.* Atlanta, GA: The Council for School Performance.

Harris, D. N. (2007). High-flying schools, student disadvantage, and the logic of NCLB. *American Journal of Education, 113,* 367-394. doi:10.1086/512737

Haycock, K. (2006). No more invisible kids. *Educational Leadership, 64*(3), 38-42. Retrieved from http://www.ascd.org/publications/educational-leadership/nov06/vol64/num03/No-More-Invisible-Kids.aspx

Heck, R. H., Larsen, T. J., & Marcoulides, G. A. (1990). Instructional leadership and school achievement: Validation of a causal model. *Educational Administration Quarterly, 26*(2), 94-125. doi:10.1177/0013161X90026002002

Hess, F. M. (2006). Accountability without angst? Public opinion and No Child Left Behind. *Harvard Educational Review, 76,* 587-610. Retrieved from http://www.frederickhess.org/5097/accountability-without-angst

Hess, F. M., & Petrilli, M. J. (2006). *No Child Left Behind.* New York, NY: Peter Lang.

Hiller, S., DiTommaso, A., & Plucker, J. (2012). The evolution of Indiana's school accountability system. *Center for Evaluation and Education Policy, 10*(5). Retrieved from http://ceep.indiana.edu/projects/PDF/PB_V10N5_2012_EPB.pdf

Hoff, D. J. (2002, October 9). States revise the meaning of 'proficient.' *Education Week, 11*(39), 9-13. Retrieved from http://www.edweek.org/ew/articles/2002/10/09/06tests.h22.html

Hopkins, D., & Ellis, P. (1991). The effective small primary school: Some significant factors. *School Organization, 11*(1), 115-122. doi:10.1080/0260136910110110

Howell, J., Miller, P., Pak, H. H., Sattler, D., Schack, T., Spery, E., Widhalm, S., & Palmquist, M. (2012). *Reliability and validity.* Ft. Collins, CO: Colorado State University, Writing@CSU. Retrieved from http://writing.colostate.edu/guides/guide.cfm?guideid=66

Howley, C. (1994). *The academic effectiveness of small-scale schooling* (An update). ERIC Digest. Retrieved from https://archive.org/stream/ERIC_ED372897#page/n0/mode/2up

Imms, T. (2008). *No child left behind: An analysis of the controversy*. Retrieved February 20, 2015, from http://www.newfoundations.com/ConAnalysis/Imms601F04.html

Improving America's School Act of 1994. (1994). *Public Law 103 – 382.* 103rd Congress. 2nd Session. 25th January, 1994.

Indiana Department of Education. (n.d.). *New elementary & middle school (E/MS) PL 221 "A-F" Model.* Retrieved May 23, 2014, from http://www.doe.in.gov/sites/default/files/accountability/basic-summary-f_1.pdf

Indiana General Assembly. (2015). *Indiana General Assembly: 2015 Top Legislative Subjects.* Retrieved January 22, 2015, from http://iga.in.gov/information/stats/legislative-subjects/

Irmsher, K. (1997). *School size.* ERIC Digest, Number 113. Retrieved from http://files.eric.ed.gov/fulltext/ED414615.pdf

Jacobs, B. (2002). Where the boys aren't: Non-cognitive skills, returns to school and the gender gap in higher education. *Economics of Education Review, 21,* 589-598. doi:10.3386/28964

Jensen, E. (2009). *Teaching with poverty in mind: What being poor does to kids' brains and what schools can do about it*. Alexandria, VA: Association for Supervision and Curriculum Development.

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

Kane, T. J., Douglas, O. S., & Geppert, J. (2002). Randomly accountable. *Education Next, 2*(1), 57-68. Retrieved from http://educationnext.org/randomly-accountable/

Kantrowitz, B. (1993, April 18). A nation still at risk. *Newsweek, 121*(16), 46-49. Retrieved from http://www.newsweek.com/nation-still-risk-191842

Keller, B. (1998). Principals matter. *Education Week on the Web.* Retrieved from http://www.edweek.org/ew/1998/11prin.h18

Keppel, F. (1966). *The necessary revolution in American education.* New York, NY: Harper & Row.

Klau, K., & Autry, W. (2010). *Key elements for educational accountability models in transition: A guide for policymakers.* Washington, DC: Council of Chief State School Officers. Retrieved from http://www.ccsso.org/Documents/2010/Key_Elements_ for_Accountability_2010.pdf

Leithwood, K., Louis, K.S., Anderson, S., & Wahlstrom, K. (2004). *Review of research: How leadership influences student learning*. New York: Wallace Foundation.

Lewis, A. C. (2006). Clean up the test mess. *Phi Delta Kappan, 87,* 643-645. doi:10.1177/003172170608700902

Ligon, G. (2008). *The optimal reference guide: Comparison of growth and value-add models growth model series – part II.* Retrieved from http://www.espsolutionsgroup.com/espweb/assets/files/ESP_Comparison_of_Growth_and_Value_Add_Models_ORG.pdf

Linn, R. L. (1999). Validity standards and principles on equity in educational testing and assessment. In A. L. Nettles & M. T. Nettles (Eds.). *Measuring up: Challenges minorities face in educational assessment* (pp. 13-33). Boston, MA: Kluwer Academic.

Linn, R. L. (2000). Assessments and accountability. *Educational Researcher, 29*(2), 4-16. doi:10.3102/0013189X029002004

Linn, R. L. (2005, September 8-9). *AYP for schools: Consequences of state accountability design decision.* Presentation to the 2005 CRESST conference, "Celebrating 20 Years of Research on Educational Assessment," University of California, Los Angeles.

Linn, R. L. (2008). Methodological issues in achieving school accountability. *Journal of Curriculum Studies, 40,* 699-711. doi:10.1080/00220270802105729

Linn, R. L., & Haug, C. (2002). Stability of school-building accountability scores and gains. *Educational Evaluation and Policy Analysis, 24*(1), 29-36. Retrieved from http://www.jstor.org/stable/3594145

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

Loeb, S., Kalogrides, D. & Horng, E. L. (2010). Principal preferences and the uneven distribution of principals across schools. *Educational Evaluation and Policy Analysis, 32*(2). Retrieved from http://files.eric.ed.gov/fulltext/ED509692.pdf

Lomax, R., & Hahs-Vaughn, D. (2012). *Statistical concepts: A second course* (4th ed.). New York, NY: Routledge.

Lowen, A., Haley, M., & Burnett, N. (2010). To consolidate or not to consolidate, that is the question: Optimal school size and teacher incentive contracts. *Academy of Educational Leadership Journal, 14*(3), 1-14.

Maleyko, G., & Gawlik, M. (2011). No Child Left Behind: What we know and what we need to know. *Education, 131,* 600-624. Retrieved from http://krboernke.files.wordpress.com/2011/05/nclb.pdf

Marion, S., White, C., Carlson, D., Erpenbach, W., Rabinowitz, A., & Sheinker, J., (2002). *Making valid and reliable decisions in determining adequate yearly progress.* Washington, DC: Council of Chief State School Officers.

Marshall, G. (1998). *Coleman report: A dictionary of sociology.* Retrieved July 24, 2013, from http://www.encyclopedia.com/doc/1O88-ColemanReport.html

Mathis, W. J. (2004). NCLB and high-stakes accountability: A cure? Or a symptom of the disease? NCLB: Failed schools--or failed law? *Educational Horizons, 82*(2), 143-152. Retrieved from http://files.eric.ed.gov/fulltext/EJ684639.pdf

Mathis, W. J. (2006). *The accuracy and effectiveness of adequate yearly progress: NCLB's school evaluation system.* East Lansing, MI: Great Lakes Center for Education Research and Practice. Retrieved from http://www.greatlakescenter.org/docs/Policy_Briefs/GLC_AYP_Mathis_FINAL.pdf

National Association of State Boards of Education. (2002). *The No Child Left Behind Act: What states need to know?* Retrieved November 17, 2005, from http://www.nasbe.org/Front_Page/NCLB/NCLBFeedback_validation.html

National Commission on Excellence in Education. (1983). *A nation at risk.* Washington, DC: Author.

No Child Left Behind Act of 2001. (2002). *Public Law 107 – 110.* 107th Congress. 2nd Session. 8th October, 2002.

Peterson, P. E., & West, M. R. (Eds.). (2003). *No child left behind: The politics and practice of school accountability.* Washington, DC: Brookings Institution Press.

Porter, A. C., Linn, L., & Trimble, C. S. (2005). The effects of state decisions about NCLB adequate yearly progress targets. *Educational Measurement: Issues and Practice, 24*(4), 32-39. Retrieved from http://kmaier.wiki.educ.msu.edu/file/view/porter+2005+effects+of+state+decisions+about+ayp+targets.pdf

Ravitch, D. (1983). *The troubled crusade: American education, 1945-1980.* New York, NY: Basic Books.

Rector, R., Johnson, K., & Fagan, P. (2001). *Understanding differences in Black and White child poverty rates* (Report No. CDA01-04 of The Heritage Center for Data Analysis). Washington, DC: The Heritage Foundation.

Riddle, W. (2005). *Adequate yearly progress (AYP): Implementation of the No Child Left Behind Act*. [Congressional Research Service Report for Congress, Order Code RL32495]. Washington, DC: Library of Congress. Retrieved from http://www.oswego.edu/~ruddy/Educational%20Policy/CRS%20Reports/NCLBA-Adequate%20Yearly%20Progress.pdf

Riddle, W. C. (2006, October 26). *Adequate yearly progress (AYP): Implementation of No Child Left Behind Act.* Washington, DC: Congressional Research Service. Retrieved May 23, 2014, from http://www.opencrs.com/rpts/RL32495_20051026.pd

Riggen, V. (2013). *School size and student achievement.* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3589555)

Ritz, G., & Yager, S. (2013, October 15). *Indiana's student-centered accountability system.* Retrieved February 8, 2015, from http://www.doe.in.gov/news/indiana-superintendent-public-instruction-glenda-ritz-and-dr-steven-yager-southwest-allen

Roe, K. D. (2006, rev. 2009). *States' impact on federal education policy project.* Albany, NY: New York State Archives. Retrieved from http://www.sifepp.nysed.gov/edindex.shtml

Rosenthal, R. (1994). Interpersonal expectancy effects: A 30-year perspective. *Current Directions in Psychological Science,* 176-179. doi:10.1111/1467-8721.ep10770698

Ryan, K. (2002). Shaping educational accountability systems. *American Journal of Evaluation, 23,* 453-468. doi:10.1177/109821400202300406

Ryan, J. E. (2004). The perverse incentives of the No Child Left Behind Act. *New York University Law Review, 79,* 932-989. doi:10.2139/ssm.476463

Sanders, W. L., & Horn, S. P. (1998). Research findings from the Tennessee's value-added assessment system (TVAAS) database: Implications for educational evaluation and research. *Journal of Personal Evaluation in Education, 12,* 247- 256. doi:10.1023/A:1008067210518

Schoen, L., & Fusarelli, L. D. (2008). Innovation, NCLB, and the fear factor: The challenge of leading 21st century schools in an era of accountability *Educational Policy, 22,* 181-203. doi:10.1177/08959040807311291

Seashore-Louis, K., Leithwood, K., Wahlstrom, K. L., & Anderson, S. E. (2010). *Learning from leadership: Investigating the links to improved student learning.* Minneapolis, MN: University of Minnesota, Center for Applied Research and Educational Improvement.

Sergiovanni, T. J. (1995). *The principalship: A reflective practice perspective* (3rd ed.). Needham Heights, MA: Allyn & Bacon.

Sizer, T. (1992). *Horace's school: Redesign of the American high school.* Boston, MA: Houghton Mifflin.

Smith, J. R., Brooks-Gunn, J., & Klebanov, P. K. (1997). Consequences of living in poverty for young children's cognitive and verbal ability and early school achievement. In G. J. Duncan & J. Brooks-Gunn (Eds.), *Consequences of growing up poor* (pp. 132-189). New York, NY: Russell Sage Foundation.

Staiger, D. O., Geppert, J., & Kane, T. J. (2002). Randomly accountable. *Education Next, 2*(1), 57-68. Retrieved from http://educationnext.org/randomly-accountable/

Sum, A. M., & Fogg, W. N. (1991). The adolescent poor and the transition to early adulthood. In P. Edeleman & J. Ladner (Eds.), *Adolescence & Poverty: Challenge for the 1990's* (pp. 37-110). Lanham, MD: Center for National Policy Press.

Superfine, B. M. (2005). The politics of accountability: The rise and fall of Goals 2000. *American Journal of Education, 112*(1), 10-43. doi:10.1086/444513

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2,* 53-55. Retrieved from http://www.ijme.net/archive/2/cronbachs-alpha.pdf

Teddlie, C., Reynolds, D., & Sammons, P. (2000). The methodology and scientific properties of school effectiveness research. In C. Teddlie & D. Reynolds (Eds.), *The international handbook of school effectiveness research* (pp. 55-133). New York, NY: Falmer Press.

Teskie, P., & Schneider, M. (1999). *The importance of leadership: The role of school principals.* Arlington, VA: The PricewaterhouseCoopers Endowment for The Business of Government.

Thompson Publishing Group. (2006). *As state plans go to peer review, some question growth decisions.* Retrieved May 17, 2014, from http://www.thompson.com/libraries/titleonline/news_desk/tio060508a.html

Thurlow, M. (2009). *Accountability.* Retrieved July 8, 2014, from http://www.education.com/reference/article/accountability/?&session-id=d0df1ef083e6182f141eb48b892288d6

Toppo, G. (2004, April 24). 'Nation at Risk': The best thing or the worst thing for education? *USA Today.* Retrieved September 10, 2014, from http://legacy.kctcs.edu/todaysnews/index.cfm?tn_date=2008-04-24#15741

Turner, E. (2013). *What effective principals do to improve instruction and increase student achievement* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3589587)

United States Department of Education. (2001). *The longitudinal evaluation of school change and performance (LESCP) in Title I schools, Volume 1: Executive summary* (Doc. No. 2001-20). Washington, DC: Office of the Deputy Secretary.

United States Department of Education. (2002). *The No Child Left Behind Act of 2001, Executive Summary.* Retrieved from http://www2.ed.gov/nclb/overview/intro/execsumm.pdf

United States Department of Education. (2009, November 2). *Archived: OVAE: School Size.* Retrieved May 22, 2014, from http://www2.ed.gov/about/offices/list/ovae/pi/hs/schoolsize.html

Vernaza, N. A. (2012). Teachers' perceptions of high-stakes accountability in Florida's Title I elementary schools. *Current Issues in Education, 15*(1), 1-12. Retrieved from http://cie.asu.edu/ojs/index.php/cieatasu/article/viewFile/870/282

Villenas, S. (2001). Latina mothers and small-town racisms: Creating narratives of dignity and moral education in North Carolina. *Anthropology and Education Quarterly, 32*(1), 3-28. Retrieved from http://www.jstor.org/stable/3196209

Waters, J. T., Marzano, R. J., & McNulty, B. (2003). *Balanced leadership: What 30 years of research tells us about the effect of leadership on student achievement.* Aurora, CO: Mid-continent Research for Education and Learning.

Watts, J. R. (2012). Being Black in small town USA. *Diverse Issues in Higher Education, 29*(6), 23. Retrieved from http://dev.southcentral.edu/commonCourseOutline/test/diversity_committe/resources/Perspectives/Being%20Black%20in%20Small%20Town%20USA.pdf

Wiley, E. W., Mathis, W. J., & Garcia, D. R. (2005). *The impact of the adequate yearly progress requirement of the federal "No Child Left Behind" Act on schools in the Great Lakes region.* East Lansing, MI: Great Lakes Center for Educational Research & Practice. Retrieved from http://greatlakescenter.org/docs/early_research/g_l_new_doc/EPSL-0505-109-EPRU_EXEC.pdf

Yu, F. (2007) *Identifying effective and ineffective schools for accountability purposes: A comparison of four generic type of accountability models.* (Doctoral dissertation). Retrieved from http://etd.lsu.edu/docs/available/etd-06112007-225430/unrestricted/Yu_dis.pdf

Zehr, M. A. (2001, October 17). After long debate, Indiana adopts plan for ranking schools. *Education Week.* Retrieved May 19, 2014, from http://www.edweek.org/ew/articles/2001/10/17/07indiana.h21.html?print=1

Zembar, M., & Blume, L. (2009). *Middle childhood development: A contextual approach*. Upper Saddle River, N.J.: Merrill/Pearson.

Zimmer, T. E. (2007). *Economies of scale in consolidation: Size and Indiana school corporations* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3307449)

APPENDIX A: INFORMED CONSENT

August 2014

A Study on the Perception of Independent Variables Used for Accountability Purposes by Indiana Teachers, Principals, and Superintendents.

You are being invited to participate in a research study about the variables used to determine school accountability in Indiana. This study is being conducted by Keith Bush under the guidance of Dr. Terry McDaniel from the Educational Leadership, Administration, and Foundations Department at Indiana State University. This study is being conducted to fulfill a dissertation requirement. The purpose of this study is to evaluate the use of social media by principals and superintendents in Indiana. A survey is being sent to all public school teachers, principals, and superintendents in the State of Indiana.

There are no known risks if you decide to participate in this research study. There are no costs to you for participating in the study. The information you provide will help school corporations determine the factors that persuade or dissuade the use of social media by public school principals and superintendents in Indiana. The questionnaire will take approximately five minutes to complete. The information collected may not benefit you directly, but the information learned in this study may provide more general benefits.

This survey is anonymous. You will not be asked to share your name at any time. This is a web-based survey, although there is no absolute guaranteed anonymity, there will be no collection of any participants' IP addresses or any attempt to identify the names of the participates by the researcher. You may delete this e-mail in which this message was delivered at anytime. There will be no future e-mail contacts concerning this survey. In addition, no one will be able to identify you or your answers, and no one will know whether or not you participated in the study. Individuals at the ISU Institutional Review Board may inspect these records. Should the data be published, no individual information will be disclosed.

Your participation in this study is voluntary and *extremely appreciated*. By completing parts or all of the survey through the Qualtrics program, you are voluntarily agreeing to participate. You are free to decline to answer any particular question you do not wish to answer for any reason. At anytime, you may close the browser and exit the program if you do not wish to complete the survey after starting the process.

If you have any questions about the study, please contact me at (812) 542-8503 or at kbush@nafcs.k12.in.us. You may also contact my faculty sponsor, Dr. Terry McDaniel, at (812) 237-3862 or at terry.mcdaniel@instate.edu.

If you have any questions about your rights as a research subject or if you feel you've been placed at risk, you may contact the Indiana State University Institutional Review Board

(IRB) by mail at Indiana State University, Office of Sponsored Programs, Terre Haute, IN, 47809, by phone at (812) 237-8217, or by e-mail at irb@indstate.edu.

APPENDIX B: PERCEPTIONS OF VARIABLES UTILIZED TO EFFECTIVELY MEASURE SCHOOL ACCOUNTABILITY

My name is Keith Bush and I am a doctoral candidate at Indiana State University. For my dissertation study, I have chosen to study the perceptions of teachers, principals and superintendents on what variables effectively measure school accountability. I am requesting your participation in this research study. Your participation is voluntary and there is no consequence if you do not participate. No one will be able to identify you as a participant. At any time you have the right to refuse to participate by simply closing the browser and exiting the program

Section I: School Accountability Variables

Directions: Please answer each question as it pertains to your opinion related to the variables utilized when measuring school accountability.

1) Were you employed as a teacher, principal, or superintendent in the state of Indiana for the 2012 -2013 school year?
   Yes No

2) I believe that student achievement results from ISTEP+ should be used to measure school accountability.
   1 = Strongly Disagree
   2 = Disagree
   3 = Somewhat Disagree
   4 = Somewhat Agree
   5 = Agree
   6 = Strongly Agree

3) I believe that it is important to look at the growth of the bottom 25% of students when determining school accountability.
   1 = Strongly Disagree
   2 = Disagree
   3 = Somewhat Disagree
   4 = Somewhat Agree
   5 = Agree
   6 = Strongly Agree

4) I believe it is important to separate the top 75% of students in an accountability model to ensure that all students are improving.
   1 = Strongly Disagree
   2 = Disagree
   3 = Somewhat Disagree
   4 = Somewhat Agree
   5 = Agree
   6 = Strongly Agree

5) I believe that schools or districts should be penalized (have points removed) for having a high percentage of students in the low-growth region.
   1 = Strongly Disagree
   2 = Disagree
   3 = Somewhat Disagree
   4 = Somewhat Agree
   5 = Agree
   6 = Strongly Agree

6) I believe that schools or districts should be rewarded (receive bonus points) for showing positive growth.
   1 = Strongly Disagree
   2 = Disagree
   3 = Somewhat Disagree
   4 = Somewhat Agree
   5 = Agree
   6 = Strongly Agree

Section II: Accountability Model Selection

7) Similar to AYP, I believe that school accountability should be measured by the percentage of students passing a standardized test for one specific year.
   1 = Strongly Disagree
   2 = Disagree
   3 = Somewhat Disagree
   4 = Somewhat Agree
   5 = Agree
   6 = Strongly Agree

8) I believe that individual student data should be observed, from one year to the next, to determine school accountability.
   1 = Strongly Disagree
   2 = Disagree
   3 = Somewhat Disagree
   4 = Somewhat Agree
   5 = Agree
   6 = Strongly Agree

9) I believe that teacher effectiveness should be included in the design of a school accountability model.
   1 = Strongly Disagree
   2 = Disagree
   3 = Somewhat Disagree
   4 = Somewhat Agree
   5 = Agree
   6 = Strongly Agree

10) I believe that an accountability model should take a student's socioeconomic level into account when measuring student performance.
   1 = Strongly Disagree
   2 = Disagree
   3 = Somewhat Disagree
   4 = Somewhat Agree
   5 = Agree
   6 = Strongly Agree

11) I believe that an accountability model should take a student's race into account when assessing student performance.
   1 = Strongly Disagree
   2 = Disagree
   3 = Somewhat Disagree
   4 = Somewhat Agree
   5 = Agree
   6 = Strongly Agree

12) I believe that an accountability model should take a student's gender into account when assessing student performance.
   1 = Strongly Disagree
   2 = Disagree
   3 = Somewhat Disagree
   4 = Somewhat Agree
   5 = Agree
   6 = Strongly Agree

Section III: Demographic information

13) What is your current position?
Teacher Principal Superintendent

14) What percentage of students in your school (teacher/principal) or school district (superintendent) received free or reduced lunch during the 2013-14 school year?
0 – 33% 34 - 67% 68-100%

15) What was the 2012-13 letter grade your school (principal) or school district (superintendent) received from the Indiana Department of Education?
A B C D or F

16) What is the best descriptor of your school (teacher/principal) or district (superintendent) locality?
Urban Suburban Rural

17) In your opinion, what variables should be utilized to measure school accountability?

18) With the fact that school accountability will be with us, how could we improve the accountability framework?

Section IV: Principal Information

Directions: This section is specific to school principals only. If you are a superintendent or a teacher, you may skip this section. Please answer each question as it pertains to your specific building.

19) What is the specific percentage of students on free or reduced lunch?

20) How many students does your school currently service?

21) How many years have you been the principal at your current school?

APPENDIX C: PERMISSION TO USE ILLUSTRATIONS

March 4, 2015

Dr. Keith Bush
Scribner Middle School
910 Old Vincennes Road
New Albany, IN 47150
Kbush@nafcs.k12.in.us

Dear Dr. Bush,

This letter is in response to your request to use the Council of Chief State School Officers (CCSSO) copyrighted material, specifically *The Policymakers' Guide to Growth Models for School Accountability: How do Accountability Models Differ?* and the *Statewide Educational Accountability under the NCLB Act* for citation in your dissertation. This letter serves to officially grant use of the aforementioned CCSSO material for print and electronic formats (nonexclusive world rights in English).

We do ask that the following formats be used when referencing the material:

> *The Policymakers' Guide to Growth Models for School Accountability: How do Accountability Models Differ?* was developed by the Council of Chief State School Officers and member states.
>
> Council of Chief State School Officers. (2005). The Policymakers' Guide to Growth Models for School Accountability: How do Accountability Models Differ? Washington, DC: Pete Goldschmidt, Pat Roschewski, Kilchan Choi, William Auty, Steve Hebbler, Rolf Blank, Andra Williams.
>
> *Statewide Educational Accountability under the NCLB Act* was developed by the Council of Chief State School Officers and member states.
>
> Council of Chief State School Officers. (2003). Statewide Educational Accountability under the NCLB Act. Washington, DC: William Erpenbach, Fore-Fast, & Potts.

CCSSO is a nonpartisan, nationwide, nonprofit organization of public officials who head departments of elementary and secondary education in the states, the District of Columbia, the Department of Defense Education Activity, and five U.S. extra-state jurisdictions. CCSSO provides leadership, advocacy, and technical assistance on major educational issues. CCSSO seeks member consensus on major educational issues and expresses their views to civic and professional organizations, federal agencies, Congress, and the public.

Thank you for your interest in this CCSSO product. We appreciate your support and interest in our organization. Please do not hesitate to contact us if we can provide any further information.

Thank you,

Melissa McGrath
Director of Communications
Council of Chief State School Officers

One Massachusetts Ave, NW • Suite 700 • Washington, DC 20001
Tel: 202.336.7000 • *Fax*: 202.408.8072 • *Web*: www.ccsso.org