

# **HOW PRINCIPALS AND TEACHERS RESPOND TO STATES' ACCOUNTABILITY SYSTEMS**

By

Hyemi Lee

A DISSERTATION

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

K-12 Educational Administration - Doctor of Philosophy

2013

## **ABSTRACT**

### **HOW PRINCIPALS AND TEACHERS RESPOND TO STATES' ACCOUNTABILITY SYSTEMS**

By

Hyemi Lee

Since the 1990s, many states have started implementing standards-based reforms and developed their own accountability systems. Under the NCLB, each state established academic content and performance standards, implemented test for all the students in third grade through eighth grade annually, and set up annual measurable objectives (AMOs) in reading and mathematics for districts, schools, and designated student subgroups within schools. The combination of states' decisions on accountability policies, such as performance standards, high school graduation exit exams, and the difference of between starting points and intermediate goals, may lead to the varying strength of the accountability systems in different states.

Although several studies focused on whether these differences are related to student achievement and teachers' instruction, little is known about how principals respond to accountability systems, although principals make a big difference in teachers' instruction and students' academic outcomes. Therefore, it may be necessary to find the relationship between the strength of the states' accountability policies and principals' responses (having influence on instruction and facilitating teachers' learning), and the relationship between the strength of states' accountability systems and teachers' responses (teacher autonomy and their participation in professional development programs).

The relationship between the strength of accountability systems (the states' proficiency performance standards, the difference of starting point and intermediate goals (AMO strength) in states, and the high school graduation exit exams) and principals' responses were studied using 2-level hierarchical linear modeling (HLM) analysis based on 2007-2008 SASS, and the relationship between the strength of accountability systems and teachers' responses were examined using 3-level hierarchical linear modeling analysis based on the same data set.

The analysis of two level HLM found the negative effects of states' accountability systems on principals responses. AMO strength was negatively related to principals' influence on instruction, and the high school graduation exit exams negatively affected principals' support of professional days before and during the school year. However, other states' accountability policies, the proficiency performance standards may not have any relationship with principals' influence on instruction and their facilitating teacher learning. Principals' professional development programs and school climate were related to principals' responses to states' accountability systems.

The findings of three level HLM showed that the proficiency performance standards increase teacher curriculum autonomy and their spending time for content professional development programs although AMOs strength and high school graduation school exit exams decreased them. Principals were an essential factor for teacher autonomy and their participation in professional development. School physical features were effective on teacher curriculum autonomy and their content professional development programs, while school climate were critical on teacher instructional autonomy and teachers' spending time in classroom management.

## ACKNOWLEDGEMENTS

I really want to express my full gratitude to my advisor and my dissertation committee chair, Dr. Susan Printy, for her assistance through my doctoral studies at Michigan State University. During four and half years, she always has trusted in me and provided academic and personal advices. She also devoted numerous hours to reading my dissertation and to providing insightful comments in order to improve my dissertation. Without her guidance, I cannot stay and develop my life at MSU.

I really appreciate my dissertation committee, Dr. BethAnn Smith, Dr. Peter Youngs, and Dr. Kristy Cooper. They provided invaluable advice and comments. Their kind words are great encouragement for me. It has been a great honor to have them serve on my dissertation committee.

I would like to extend thanks my friends. America life as a stranger was sometimes lonely and difficult, but my friends were a wonderful boost for my long years of American life. They always provide power that I can overcome the difficult time, and they make me smile. Moreover, Rose Cooper is another sister in America. Thanks to her, I can have unforgettable memory in my life.

Finally, my family has been an unfailing source of my understanding and encouragement. My parents and younger brother have offered steady faith that I would reach my goals. They have maintained continuing interest in my progress and cheered each success, and they have given unconditional support to achieve my goals. My successful life may be due to my family unwavering support. I love you.

# TABLE OF CONTENTS

LIST OF TABLES .....	viii
LIST OF FIGURES .....	x
<b>CHAPTER ONE .....</b>	<b>1</b>
<b>I. INTRODUCTION.....</b>	<b>1</b>
<b>CHAPTER TWO .....</b>	<b>9</b>
<b>II. LITERATURE REVIEW .....</b>	<b>9</b>
1. Accountability in America .....	9
1) Definition .....	9
2) Assumptions.....	10
3) History .....	11
4) Federal accountability framework .....	14
5) Changes of states' accountability systems .....	17
2. Indexes of Accountability Systems.....	21
1) Indexes of accountability systems in previous studies .....	21
2) Major factors of states' accountability strength.....	29
Proficiency performance standards .....	29
Annual measurable objectives strength .....	31
High school graduation exit exams .....	34
3. Studies of Accountability Effects .....	36
1) The effects of accountability on students .....	36
2) The effects of accountability on teachers .....	39
4. Principals' Response to Accountability .....	43
1) Having influence on instruction .....	43
2) Facilitating teacher learning .....	44
5. Teachers' Responses in Accountability.....	47
1) Teacher autonomy.....	47
2) Teachers' professional development .....	49
<b>CHAPTER THREE .....</b>	<b>52</b>
<b>III.METHODLOGY.....</b>	<b>52</b>
1. Conceptual Model.....	52
2. Research Questions and Hypotheses .....	54
3. Data .....	56

4. Variables .....	59
1) The strength of states' accountability systems .....	59
2) Principals' responses.....	61
3) Teachers' responses.....	62
4) Control variables.....	64
5. Analysis of Principal's Responses to Accountability Policies .....	67
6. Analysis of Teachers' Responses to Accountability Policies .....	72
7. Limitations .....	81
<b>CHAPTER FOUR.....</b>	<b>82</b>
<b>IV. RESULTS .....</b>	<b>82</b>
1. Principals' Responses to States' Accountability Systems.....	82
1) The level and characteristics of principals' responses.....	82
2) The relationship between the strength of states' accountability systems and principals' responses .....	85
Principals' influence on instruction .....	85
Support of professional work .....	88
Provision of professional days before or during the school year .....	90
Synthesis of principals' responses .....	92
2. Results of Teachers' Response to States' Accountability Systems .....	94
1) The level and characteristics of teachers' responses .....	94
2) The relationship between states' accountability strength and teacher autonomy .....	98
Teacher curriculum autonomy .....	98
Teacher instructional autonomy.....	102
Synthesis of teacher autonomy .....	104
3) The relationship between states' accountability strength and teachers' participation time in professional development .....	106
Content professional development participation time .....	106
Instruction professional development participation time .....	110
Classroom management professional development participation time .....	112
Synthesis of teachers' participation time in professional development.....	115
<b>CHAPTER FIVE.....</b>	<b>117</b>
<b>V. DISCUSSION, IMPLICATIONS, AND CONCLUSION.....</b>	<b>117</b>
1. Discussion .....	117
1) The weak negative relationship between states' accountability policies and principals' responses.....	117
2) The directly opposed effects of states' accountability policies on	

teachers' responses.....	120
3) The limited effects of states' accountability policies on specific schools .....	122
4) The limited effects of states' accountability policies on specific domain of practice .....	123
5) Principals' effects on teachers' responses.....	125
2. Implications.....	127
3. Conclusion .....	130
<b>APPENDICES .....</b>	<b>131</b>
Appendix A Proficiency Performance Standards in Fourth and Eighth Grade in Reading .....	133
Appendix B Proficiency Performance Standards in Fourth and Eighth Grade in Mathematics.....	134
Appendix C Starting Points of 50 States in 2002 .....	135
Appendix D Intermediate Goals of 50 States in 2007 .....	136
Appendix E Number of High School Teachers Among 50 States .....	137
Appendix F Principals Responses by States .....	139
Appendix G Teachers' Responses by States .....	140
<b>BIBLIOGRAPHY .....</b>	<b>141</b>

## LIST OF TABLES

Table II-1 Ten Principles for Accountability .....	15
Table II-2 Types of High School Graduation Exit Exams .....	19
Table II-3 Existing Indexes of Accountability .....	27
Table III-1 The Characteristics of Principal Data Set .....	57
Table III-2 The Characteristics of Teacher Data Set .....	58
Table III-3 The Characteristics of Teacher Data Set .....	58
Table III-4 The Strength of States' Accountability Systems .....	60
Table III-5 Variables of Principals' and Teachers' responses .....	63
Table III-6 Control Variables .....	66
Table III-7 Descriptive Statistics for the 2-level Analysis Variables .....	67
Table III-8 Descriptive Statistics for the 3-level Analysis Variables .....	74
Table IV-1 The Level of Characteristics of Principals' Responses .....	84
Table IV-2 The Influential Factors for Principals' Influence on Instruction .....	86
Table IV-3 The Influential Factors for Principals' Facilitating Teacher Learning .....	91
Table IV-4 The Level of the Characteristics of Teachers' Responses .....	97
Table IV-5 Influential Factors for Teacher Curriculum Autonomy .....	100
Table IV-6 Influential Factors for Teacher Instructional Autonomy .....	103
Table IV-7 Influential Factors for Teacher's Participation Time in Content Professional Development Programs .....	108
Table IV-8 Influential Factors for Teacher's Participation Time in Instruction Professional Development Programs .....	111
Table IV-9 Influential Factors for Teacher's Participation Time in Professional Development Related to Classroom Management .....	114



Table VI-1 Proficiency Performance Standards in Fourth and Eighth Grade in Reading .....	133
Table VI-2 Proficiency Performance Standards in Fourth and Eighth Grade in Mathematics...	134
Table VI-3 Starting Points of 50 States in 2002.....	135
Table VI-4 Intermediate Goals of 50 States in 2007.....	136
Table VI-5 Number of High School Teachers Among 50 States .....	137
Table VI-6 Principals Responses by States.....	139
Table VI-7 Teachers' Responses by States.....	140

## **LIST OF FIGURES**

Figure II-1 Florida Annual Measurable Objective for Reading.....	33
Figure II-2 Michigan Annual Measurable Objectives for Reading .....	33
Figure III-1 A Conceptual Model.....	53

# **CHAPTER ONE**

## **I. INTRODUCTION**

Since the Improving America's Schools Act (IASA) and Goals 2000 were established in 1994, many states have started implementing standards-based reforms. Goals 2000, the first policy based on the standards-based approach, offered states federal funding for using proficiency performance standards (McDonnell, 2005). Under the law, schools that received Title I funds developed academic standards and prepared assessment systems for measuring students' academic performance (Finn & Kanstoroom, 2001). Also, under the IASA, many states established their standards-based reforms. The law stated that schools that received Title I funding should use the state's content standards, and students in the schools should acquire the standards that states established (McDonnell, 2005). Due to the effects of IASA and Goals 2000, the number of states having accountability systems has increased (Meyer, Orlofsky, Skinner, & Spicer, 2002; Goertz & Duffy, 2001). Based on these previous educational accountability policies, President George W. Bush signed The No Child Left Behind (NCLB) act into law in January 2002.

Accountability that has dominated American education since the 1980s, assumes that a school is responsible for students' achievement, that teachers do their job to gain rewards and to avoid sanctions, and that teachers' efforts can improve students' academic outcomes. Based on these assumptions of accountability, NCLB requires each state to establish academic content and performance standards, to test all the students in grades 3 through 8 annually, to set up annual measurable objectives (AMOs) in reading and mathematics for districts, schools, and designated student subgroups within schools, and to offer rewards or sanctions based on whether or not districts and schools achieve adequate yearly progress (AYP) (Erpenbach, 2011; Taylor, Stecher,

O'Day, Naftel, & Le Floch, 2010; Le Floch et al., 2007; Forte & Erpenbach, 2006; Fast & Erpenbach, 2004; Erpenbach, Forte-Fast, & Potts, 2003).

Under NCLB conditions, the fifty states have produced various accountability systems based on their educational conditions, such as a proportion of minority students and a state size. For example, states with a higher proportion of minority students, with the greatest degrees of poverty, and with large size cities tend to have state-level tests and high school graduation exit exams (Shuster, 2012; Wei, 2012; Nichols, Glass, & Berliner, 2006; Amrein & Berliner, 2002; Carnoy & Loeb, 2002). Also, states with a high family income tend to have difficult math proficiency standards and states with a large minority population and high family income are more likely to build up ambitious annual measurable objectives for math (Wei, 2012), so students in these states should acquire high test-scores to pass the standards. However, other states do not have these things. The combination of states' decisions on accountability policies, such as performance standards, high school graduation exit exams, and the difference of between starting points and intermediate goals, leads to the varying strength of the accountability systems in different states.

Because few states implemented high-stakes tests and rigorous sanctions in the early age of NCLB, states that have these policies were considered by the researchers to be "high stakes" states (Nichols et al., 2006; M. Clarke et al., 2003; Pedulla et al., 2003; Amrein & Berliner, 2002; Carnoy & Loeb, 2002). For examples, Texas, North Carolina, and New York were high stakes states, while Iowa, New Hampshire, and North Dakota were not. However, as time goes, almost states have executed high stakes tests, and thus high school graduation exit exams and big differences between their AYP starting points and their intermediate goals can be considered as a critical indicator of strong accountability systems (Wei, 2012).

After accountability systems became widespread throughout America, many researchers began studying the effects of accountability on students, e.g., whether or not accountability policies enhance students' academic accomplishments and reduce achievement gaps. Some obtained negative findings: the policies decreased reading achievement and did not reduce the achievement gap (J. Lee & Reeves, 2012; Usher, 2012; Schneider, 2011; J. Lee, 2006) although others found positive effects of accountability policies (Dee & Jacob, 2011; Reback, Rockoff, & Schwartz, 2011; Ladd & Lauen, 2010).

In addition, researchers focusing on the effects of accountability on teachers produced not only positive results but also negative results. Advocates insist that accountability systems encourage teachers to align standards and instruction with tests (Hamilton, Stecher, Russell, Marsh, & Miles, 2008; Finnigan & Gross, 2007) and to collaborate with each other (Diamond, 2007). However, opponents found that accountability narrows the curriculum, emphasizes teaching for tests (Diamond, 2012; Cocke, Buckley, & Scott, 2011; Srikantaiah, 2009), and increases teachers' stress and turnover (Hannaway & Hamilton, 2008).

Moreover, because each state has a different accountability system, several studies focused on whether these differences of states' accountability systems are related to student achievement and teachers' instruction. Some found that accountability strength is significantly related to high mathematics attainment for fourth grade Hispanic students and eighth grade African American students (Wei, 2012; Nichols et al., 2006; Carnoy & Loeb, 2002). However, others did not find any positive effects of strong stakes accountability policies (Amrein & Berliner, 2002). Several educators discovered that teachers in high-stakes states implement instruction focusing on tests more so than do those in low-stakes states (M. Clarke et al., 2003; Pedulla et al., 2003; Swanson & Stevenson, 2002).

However, little is known about how principals respond to accountability systems (Rice, 2010; McGhee & Nelson, 2005), although principals can make a big difference in teachers' instruction and students' academic outcomes (Louis, Leithwood, Anderson, & Wahlstrom, 2010; DeMoss, 2002). The existing studies about principals in the era of accountability have focused on principals' desirable responses to accountability (Elmore, 2005) and their perception of accountability policies (McKay, 2011; Kelley, Kimball, & Conley, 2000). The variable conditions described above may provide an opportunity to examine how principals respond in order to meet the goals of accountability policies and to increase students' academic achievements.

Little is known about differences in principals' responses in states with strong policies versus states with weak policies. As differences of strength in states' accountability policies make a difference in students' performance (Wei, 2012; Carnoy & Loeb, 2002) and in teachers' instruction (M. Clarke et al., 2003; Pedulla et al., 2003; Swanson & Stevenson, 2002), it is reasonable to think that the differences of states' accountability systems may also influence principals' responses. When states are more influential in developing standards for curriculum, student performance, and assessment, schools may be more accountable for student outcomes (Fuhrman & Elmore, 2004) and may experience huge stress, which can influence how they lead others (Knobl, 2010; Priolo, 2010). Due to force from states' accountability policies, principals in strong states' accountability systems may focus on methods for increasing students' academic performance than those in weak accountability systems.

Existing studies have focused on principals in states with a long history and/or strength of accountability policies. Principals in these states, including Florida, Maryland, New Jersey, and Virginia, tend to focus on students' performance (Hamilton et al., 2007), to emphasize instruction

through evaluating teachers (Gonzalez, 2012; Rutledge, Harris, & Ingle, 2010), and to establish a school environment for supportive of professional capacity (Sanzo, Sherman, & Clayton, 2011; Arbogast, 2004). A solid research base in states with variable accountability conditions is, however, non-existent. We do not know whether principals in states moderate or weak accountability systems respond identically to accountability policies as those counterparts in the strong accountability states. Therefore, I would like to study the relationship between the states' accountability systems and principals' responses.

I will specifically focus on principals' two responses: having influence on instruction and facilitating teachers' learning. In accountability contexts, students' academic outcomes are considered as a main indicator of school education success or failure (Foy, 2008). When students' test scores are not high enough to pass states' performance standards, schools and principals may receive sanctions (Mintrop & Sunderman, 2009). To avoid sanctions, principals should make efforts to increase students' academic achievements. The representative methods that principals can take to improve students' academic outcomes are an emphasis on standards and curricula, an evaluation of teachers, and an encouragement of teachers' professional development participation (Bottoms, 2003).

To acquire high-test scores, principals try to align schools' standards and curriculum with the state's standards or assessments (Hamilton et al., 2007), and they also observe in classrooms and evaluate teachers' instruction to check whether or not teachers implement schools' standards and curriculum (Gonzalez, 2012; Louis et al., 2010). Because teachers' instruction that is highly related to students' outcomes has been considered as a major issue since the emergence of accountability policies (Sebastian & Allensworth, 2012), principals may have more influence for developing teachers' capacities from professional development (Rutledge et al., 2010; Hill, 2007).

Therefore, I would like to address the first research question: what the relationship between strength of states' accountability systems and principals' responses (their influence on instruction and facilitation of teachers' learning) is. Because principals can be influenced by each state accountability system, their responses to accountability policies may not be uniform. I assume that state's high proficiency performance standards, AMO strength, and high school graduation exit exams will be significantly positive correlated with principals' influence on instruction and their facilitation of teachers' learning. Principals in states with high achievement goals and high school graduation exit exams may have more influence on standards, curriculum, and instruction, and that they facilitate teachers' learning than principals in states with weak accountability systems.

In addition, I address the second research question: what the relationship between strength of states' accountability systems and teachers' responses (teacher autonomy and their participation in professional development programs) is. I assume that states' high proficiency performance standards, AMO strength, and high school graduation exit exams will be negatively and significantly correlated with teacher autonomy and that states' high proficiency performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with teachers' participation time in professional development programs. Teachers in states with high proficiency performance standards, big differences between starting points and annual objectives, and high school graduation exit exams may produce lower level of teacher autonomy and participate in more professional development programs than teachers in states with weak accountability systems.

Especially, I assume that different principals' responses may influence the relationship between states' accountability strength and teachers' responses, such as teacher autonomy and



professional development participation time. Principals are likely to implement accountability policies in their schools, so they may influence teachers in their schools. When principals have more tight and direct power about curriculum and instruction, teachers may have low control (Eden, 2001). In addition, as a builder, designer, and supporter of professional development, principals promote teachers' participation in professional development programs (Sanzo et al., 2011; Wahlstrom & York-Barr, 2011).

To respond to these research questions, in the Chapter 2, I first will explain conception, assumption, and history of accountability. I will also describe the maturation of and changes in federal and state accountability policies since the implementation of NCLB and examine the research studying the strength of accountability systems and the studies related to accountability effects on students and teachers. Finally, I will investigate the strength of accountability systems using (1) the states' proficiency performance standards, (2) AMO strength (the difference of starting point and intermediate goals in states), and (3) high school graduation exit exams based on states' Consolidated Application Accountability Workbook.

In the Chapter 3, I will describe a conceptual map, research questions, data sets, variables, analysis, and limitations. In the Chapter 4, I will try to respond research questions. First, I will study the relationship between the strength of accountability systems and principals' responses: principals' influence on instruction and their support of professional development using 2-level hierarchical linear modeling analysis based on 2007-2008 SASS. Next, I will examine the relationship between the strength of accountability systems and teachers' responses: teacher autonomy and their participation time in professional development using 3-level hierarchical linear modeling analysis based on the same data set.

My study is intended to expand my understanding of the potential influence of

accountability policies. From this study, I can confirm dissimilar states' accountability systems. The federal accountability policies do not offer specific regulations. Under the ambiguity, each state should create and implement its accountability systems, including academic content standards, proficiency performance standards, measurement methods, assessment systems, and rewards or sanctions for schools. The combination of these factors can produce different level of states accountability systems.

In addition, I can comprehend the relationship between states' accountability systems and principals' and teachers' responses. Existing studies have focused on states with high states' accountability systems and have studied how principals and teachers respond to these states' accountability systems. However, as we know the different level of states' accountability systems, principals and teachers may differently behave based on their states' accountability systems.

## **CHAPTER TWO**

### **II. LITERATURE REVIEW**

This chapter will explain literature reviews related to accountability. To understand accountability, the definition, assumptions, and history of accountability are described. Based on the basic knowledge about accountability, this chapter will elucidate how federal government and states' implement accountability systems. In addition, indexes of accountability systems in previous studies and the effects of accountability on students and teachers are expounded. Finally, this chapter can show how principals and teachers respond to states' accountability systems in previous studies.

#### **1. Accountability in America**

This part will expound definitions, assumption, and history of accountability. Additionally, federal accountability framework and states' accountability systems are going to be explained.

##### **1) Definition**

Even prior to the federal government established the NCLB Act, the concept of accountability became prevalent. Literally, accountability comes from the verb "account" which means "to reckon, count, count up or calculate" (Wagner, 1989, p. 7). In the concept of accountability, there are at least two actors: "those being called into account; and those doing the calling" (Walberg, 2002, p. 157), and there are two factors: responsibility and entitlement (Wagner, 1989). One actor has the responsibility for giving an account, and this responsibility

comes from the law or people sharing this responsibility (Leithwood & Earl, 2000). The other actor has entitlement to demand an account. Applying this concept to education, a school has the responsibility for establishing educational goals, pursuing them, and choosing instructional methods. Parents may be entitled to ask about their children's education and school life under the law, and citizens and taxpayers are entitled to inquire about expenditures of school funds (Wagner, 1989).

Based on this concept, Rothman (1995) defined educational accountability as "the process[es] by which school districts and states attempt to ensure that schools and school systems meet their goals"(p. 189). Educational accountability policies are methods for states or school districts to check whether or not a school meets the state's educational goals.

## **2) Assumptions**

Accountability is based on several assumptions (Kozar, 2011; Ladd, 1996). The first is that the school is a basic unit delivering education and thus teachers and principals should be held accountable. The second assumption is that schools are responsible for students' performance. The third assumption is that students' academic outcomes are measured by tests and standards created by external organizations created. The final assumption is that the students' academic results become a standard to reward successful schools or to punish unsuccessful schools. In addition, accountability assumes that to gain rewards and to avoid sanctions, school staffs will do a better job of improving students' academic achievements (Finnigan & Gross, 2007; Spillane, Diamond, Burch, & Hallett, 2002). In fact, we assume both that accountability policies are effective means to influence schools and that schools have the capacities to locate, select, and implement effective improvement programs and policies for achieving accountability

(Gross & Goertz, 2005).

### **3) History**

Accountability reforms are not new in the education field: the concept of accountability has continued since 1950s. Linn (2000) mentioned five waves of reforms from 1950s to 1990s in the America. They are:

1950s: Tracking and selection

1960s: Program accountability

1970s: Minimum competency testing

1980s: School and district accountability based on standardized tests

1990s: Standards based accountability systems.

The emphasis on accountability started from the late 1950s. When the Soviet Union succeeded the Sputnik Launch, it was believed that America education “was too sluggish to respond promptly to the new demands or to make good use of science and technology for the engineering of change” (Chase, 1971, p. 182). Public education became accountable for nation priority. As accountability models in education, tests were considered as important tools for selecting students for higher education (Linn, 2000).

Through the Coleman (1966)’ report, the Equality of Educational Opportunity, more commonly known as the “Coleman Report,” educators found that students have different educational opportunities and resources based on their race and social economic status. To reduce these differences, the federal government initiated the Elementary and Secondary Education Act (ESEA) in 1969 (Linn, 2000). Under the Title I of the ESEA, the federal government spent federal funding on educational programs that are expected to improve students’ academic outcomes and the government wanted to evaluate effectiveness of these programs

using measured outcomes (Shepard, 2008). The focus of educational evaluation shifted from inputs or resources to outputs or results by title I (Ravitch, 2002).

In the 1970s, minimum competency testing reforms were widespread. In 1969, the Education Commission of the States created the National Assessment of Educational Progress (NAEP) to "examine achievement in ten learning areas, to spot changes in the level of achievement over the years and to apply the implication of those changes to national educational policy" (Wise, 1979, p. 9). With the NAEP, the number of states implementing minimum competency testing increased from 2 to 34 during ten years (Linn, 2000). Especially, states used this testing as requirement of high school graduation because this testing can check students' basic skills and evaluate public schools (Resnick, 1980).

Although accountability remained a significant topic in the 1970s, the introduction of the A Nation at Risk report in 1983 by the National Commission on Excellence in Education encouraged national awareness about accountability. The report considered public education as a main reason of the ineffective nation (Education, 1983). Since the release of A Nation at Risk report, states and the government implemented standardized test (Linn, 2000) and had more influence on school reform and more enhanced educational standards (D. L. Stevenson & Schiller, 1999; Fuhrman, Clune, & Elmore, 1988). In the 1980s, 275 state-level educational reforms were established (Wirt & Kirst, 1989, pp. 3-4). This trend continued throughout the 1990s.

In the 1990s, the federal government has encouraged states to establish and develop content and performance standards under the Improving America's Schools Act of 1994 (IASA) and Goals 2000. Under the IASA and Goals 2000, states establish challenging standards, implement assessment systems for measuring students' academic performance, and hold schools

accountable for all students' achievement (McDonnell, 2005; Finn & Kanstoroom, 2001; Goertz, 2001). Due to the effects of IASA and Goals 2000, the number of states having accountability systems increased (Meyer et al., 2002; Goertz & Duffy, 2001), but the states did not yet have completed state level accountability systems.

In 2002 the U.S. Congress passed the No Child Left Behind Act (NCLB), reauthorization of the Elementary and Secondary Education Act. Under the NCLB, each state should establish and develop mandatory national accountability systems that held schools and districts responsible for student achievement (Taylor et al., 2010; Le Floch et al., 2007). NCLB requires schools received federal Title I funding to meet their state's performance standards or to receive sanctions (Erpenbach et al., 2003). However, from 2011, the federal government has offered states the opportunity to waive several requirements of NCLB. As of March 2013, 48 states, the District of Columbia, Puerto Rico, and the Bureau of Indian Education have received waive application permit.

One current accountability policy is Race to the Top (RTT) announced by President Barack Obama in 2009. RTT was designed to produce effective school reforms by relying on incentives, not sanctions, so states that have demonstrated students' academic development and have rigorous reforms receive federal educational funds (McGuinn, 2012; G. A. Scott, 2011). RTT requires several criteria that states should establish to apply for RTT funds and these requirements led to school reform in state-level (M. McNeil, 2011). Under the RTT, forty-eight states have signed on to the Common Core State Standards Initiatives (Finn, 2012; Ravitch, 2010).

To sum up, for seven decades federal government has established various educational accountability policies, and these policies have moved from input accountability focusing on

regulations to outputs accountability focusing on students' test scores and graduation rates (Goertz, 2001; Fuhrman, 1999; Elmore, Ableman, & Fuhrman, 1996). With a tendency of federal accountability policies, states' educational accountability systems have also emphasized educational outcomes (Crowe, 2011) and have narrowed educational attention that federal government advocated (McGuinn, 2012).

#### **4) Federal accountability framework**

The most recent federal accountability in education is NCLB. In this part, I will explain major features of NCLB. The NCLB Act requires each state to design and implement its accountability systems based on ten criteria that are known as “the ten principles for accountability”. States describe academic standards, assessment systems, AYP (Adequate Yearly Progress), and rewards and sanctions in a Consolidated Application Accountability Workbook<sup>1</sup>. Ten principles for accountability are explained in Table II-1.

First, NCLB requests that states set up challenging academic content and performance standards (NCLB, 2001 sec. 1111 (b) (1)). Content standards explain what students in elementary and secondary school must know and be able to do, contain coherent and rigorous content, and encourage the teaching of advanced skills. These standards are applied to all schools and children in the state. States should establish content standards at least in mathematics, reading or language arts, and science (beginning in the 2005–2006 school year).

---

<sup>1</sup> All the state's accountability workbooks are listed on the Department of Education website (<http://www.ed.gov/leaders/leadlaccount/stateplans03/index.html>).



**Table II-1 Ten Principles for Accountability**

---

i.	A single statewide Accountability System is applied to all public schools and LEAs (local educational agencies—commonly referred to as “school districts”);
ii.	All students are included in the State Accountability System;
iii.	State definition of AYP (adequate yearly progress) is based on expectations for growth in student achievement that is continuous and substantial, such that all students are proficient in reading or language arts and mathematics no later than 2013–2014;
iv.	State makes annual decisions about the achievement of all public schools and LEAs;
v.	All public schools and LEAs are held accountable for the achievement of individual student groups;
vi.	State definition of AYP is based primarily on the state’s academic assessments;
vii.	State definition of AYP includes graduation rates for public high schools and an additional indicator selected by the state for public middle and public elementary schools (such as attendance rates);
viii.	AYP is based on reading or language arts and mathematics achievement objectives;
ix.	State Accountability System is statistically valid and reliable; and
x.	In order for a public school or LEA to make AYP, the state ensures that it assessed at least 95 percent of the students enrolled in each student group.

---

NCLB also describes performance standards, which determine how well children are mastering the material in the states’ academic content standards. Based on the degree to which students understand and master content standards, performance standards are classified into three levels: basic, advanced, and proficient. When students master the academic materials, the students are placed in the proficient level. However, basic level is the third level of achievement in providing complete information about the progress of the lower-achieving children toward mastering the proficient and advanced levels of achievement.

In addition, NCLB asks that the implementing state’s academic assessments review the annual progress of each school (NCLB, 2001 sec. 1111 (b) (3)). During the 2002-2003 school year, NCLB required reading and mathematics tests for students in three grade spans (3-5, 6-9, 10). The 2005-2006 school year increased reading and mathematics tests for all students in grades 3-8 and one grade in grade 10-12. From the 2007-2008 school years, students were required to take science tests. The assessment has to be aligned with the state’s challenging

academic content and performance standards and has to provide coherent, valid, and reliable information about student attainment of such standards. The students' achievement is disaggregated by ethnicity, gender, English proficiency, disability status, migrant status, and economic status.

Moreover, NCLB includes other academic performance indicators (NCLB, 2001 sec. 1111 (b) (2)). For example, student attendance, retention rate, state or district level assessments, and percentage of students completing special programs (advanced placement courses, gifted programs, or college preparatory courses) can be indicators (Mills, 2008). In secondary schools, the graduation rate is an indispensable indicator.

Based on these test scores and indicators, states "identify for school's improvement" whether the school makes AYP as defined in the State's plan (NCLB, 2001 sec. 111 (b) (2) (C)). To evaluate AYP, each state establishes a starting point based on the 2001- 2002 school year and a timeline for all students in each group to meet and exceed the proficient level of academic achievement by the 2013- 2014 school year. Also, states build annual measurable objectives (AMOs) and intermediate goals of assessment and other indicators in order to meet 100% proficiency by the 2013 - 2014 school year. In analyzing AMOs, no less than 95% of the students enrolled in a school must participate in the assessment programs, because when the number of students who participate in the assessment is too small, the reliability and validity of the AMOs may be damaged.

Also, NCLB explains the rewards and sanctions when schools and districts pass or fail AYP standards (NCLB, 2001 sec. 1116 (b)). When schools approach the standards of AYP or are highly ranked in its accountability systems, they gain a title of "distinguished schools" or "Honor School of Excellence". However, if schools that received Title1 funding fail to reach AYP

standards for two consecutive years, the schools are placed in “program improvement” status. In the first year of having program improvement status, failed schools can receive supplemental educational services and technical assistance. Simultaneously, they provide notice of their AYP failure to parents, and they offer opportunities for students to transfer to another public school. When schools in the improvement status category do not show any development, the schools’ staff is replaced and the schools may be reorganized or closed.

In summary, the federal government does not provide specific regulations and encourages states to create and implement their accountability systems involving these components: academic content and performance standards, measurement and assessment systems, and rewards or sanctions for schools.

### **5) Changes of states’ accountability systems**

Although 50 states, the District of Columbia and Puerto Rico received approval for their first accountability plans in June 2003, they have modified and developed their plans annually. Individual states have negotiated their educational accountability systems with the federal government in order to gain flexibility in implementing the systems (Mills, 2008) and to temporarily reduce the number of schools labeled as failing (Sunderman, 2006). In every year from 2003 to 2011, most states wanted to modify their accountability systems. 47 states in 2003-2004 school years, 20 states in 2004-2005 school year, 48 states in 2005-2006 school year, 49 states in 2008-2009, 36 states in 2009-2010, and 31 States in 2010-2011 requested modification of their accountability systems (Erpenbach, 2011; Taylor et al., 2010; Le Floch et al., 2007; Forte & Erpenbach, 2006; Fast & Erpenbach, 2004; Erpenbach et al., 2003).

One of the changes was shown in content standards. In the early stage of NCLB, there

were some variations in the content standards of what students in elementary and secondary must know (Finn & Kanstoroom, 2001). However, after the Kindergarten-12 Common Core State Standards in English and mathematics was created by the National Governors Association Center for Best Practices and the Council of Chief State School Officers in 2010, as of July 2012, forty-five states and three territories have accepted these standards. Fifty states have similar levels of content standards.

Second, performance standards also have changed. According to the studies the National Center for Education Statistics (NCES), states' performance standards for fourth and eighth grade reading were arranged from below the NAEP basic level to below the NAEP proficient level from 2003 to 2009 and standards for fourth and eighth grade mathematics have placed in little higher position than those for reading since 2003 (see Appendix A and Appendix B) (National Center for Education Statistics, 2011; Bandeira de Mello, Blankenship, & McLaughlin, 2009; National Center for Education Statistics, 2007).

Since 2003 Massachusetts has continued high standards in fourth and eighth grade reading and Mathematics since 2003 although Tennessee and Georgia have continuously had low standards. Also, some states have increased their performance standards, but other states have decreased. For example, Indiana, North Carolina, and Oklahoma increased fourth and eighth grade reading and mathematics performance standards; however, Maine, South Carolina, and Wyoming's reading and mathematics performance standards decreased (National Center for Education Statistics, 2011; Bandeira de Mello et al., 2009; National Center for Education Statistics, 2007; Peterson & Hess, 2005).

Third, there were changes in assessment fields. States have created new assessments or modified existing assessments for reading and mathematics since 2005 and for science since

2007. The number of states also increased, so in 2006-2007, thirty five states used attendance rate as other academic indicators in elementary and middle schools (Taylor et al., 2010). In addition, the number of states that implemented high school graduation exit exams has increased as shown Table II-2 below. Although nineteen states implemented graduation tests in 2002, 26 states used high school graduation exit exams in 2012 (McIntosh, 2012; Chudowsky, Kober, Gayler, & Hamilton, 2002). Especially, the number of states that implement end-of-course exams as graduation tests has increased (Zabala, Minnici, McMurrer, & Briggs, 2008).

**Table II-2 Types of High School Graduation Exit Exams**

Year	High school graduation exit exams		No mandatory exit exam
	Comprehensive exams	End-of-course exams	
2002	AL, FL, GA, IN, LA, MD, MN, MS, NV, NC, NJ, NM, OH, SC, TN, TX, VA	NY, TX <sup>2</sup>	AK, AR, AZ, CA, CO, CT, DE, DC, HI, ID, IL, IA, KS, KY, MA, ME, MI, MO, MT, NE, NH, ND, OK, OR, PA, RI, SD, UT, VT, WA, WV, WI, WY
2008	AK, AL, AZ, CA, FL, GA, ID, IN, LA, MA, MN, NC, NJ, NM, NV, OH, SC, TX, WA	MS, NY, TN, VA	AR, CO, CT, DE, DC, HI, IL, IA, KS, KY, MD, ME, MI, MO, MT, NE, NH, ND, OK, OR, PA, RI, SD, UT, VT, WV, WI, WY
2012	AL, AR, AZ, CA, FL, GA, ID, MA, MN, NJ, NV, NM, OH, OR, RI, SC, TX, WA	AK, IN, LA, MD, MS, NY, OK, VA	CO, CT, DE, DC, HI, IL, IA, KS, KY, ME, MI, MO, MT, NC, NE, NH, ND, PA, SD, TN, UT, VT, WV, WI, WY

Even though development of states' accountability systems, NCLB faces difficulties reaching its goals. For example, it may be an unachievable goal for all students in each group to meet and exceed the proficient level of academic achievement by the 2013- 2014 school year (Shelly, 2012). In 2011, to help alleviate this unattainable, the federal Department of Education received waive applications for changing their own accountability systems. States can receive

---

<sup>2</sup> Texas implemented the Texas Assessment of Academic Skills (TAAS) test and end-of-course exams (Chudowsky et al., 2002).

flexibility several aspects, such as reconfiguration of performance proficiency standards, assessment of students' academic outcomes, and identification of low-performing schools (K. S. Berry & Herrington, 2011). In addition, states should implement four requirements to obtain flexibility: “adopting college- and career-ready standards; creating state-defined accountability systems that reward success and promote improvement; strengthening teacher and principal practice through evaluation systems, and reducing duplication and administrative burden placed on districts and schools (Ayers, 2011, p. 6)”.

As of March 2013, 48 states, the District of Columbia, Puerto Rico, and the Bureau of Indian Education have submitted requests for flexibility<sup>3</sup>. Of those waiver requests, 35 have been approved and fourteen<sup>4</sup> waivers are still under review. However the California's request was rejected, and just two states, Montana and Nebraska, have not submitted applications.

---

<sup>3</sup> All information are taken from the website (<http://www.ed.gov/esea/flexibility/requests>).

<sup>4</sup> Alabama, Alaska, the Bureau of Indian Education, Hawaii, Illinois, Iowa, Maine, New Hampshire, North Dakota, Pennsylvania, Puerto Rico, Texas, Vermont, West Virginia, and Wyoming.

## **2. Indexes of Accountability Systems**

This part will introduce indexes of accountability systems that previous studies implemented and three major factors for accountability strength: the proficiency performance standards, the annual measurable objectives (AMOs) strength, and high school graduation exit exams.

### **1) Indexes of accountability systems in previous studies**

According to my analysis of federal and individual states' educational accountability policies, each state has different accountability systems (McDermott, 2003). The NCLB law does not mention specific accountability systems; so each state, in different ways, interprets, designs, implements, and develops its own accountability policies (Heinecke, Curry-Corcoran, & Moon, 2003). Academic content standards, performance standards, assessment systems, AYP, and AMOs vary substantially among states. Also, the states use different “rewards, sanctions, selection criteria for low-performing schools, exit criteria for probation, school governance requirements, planning mandates, monitoring systems, and supports for building capacity at schools” (Mintrop, 2003, p. 3).

These differences in accountability policies among the 50 states create different accountability strength. The Council of Chief State School Officers (CCSSO) insists that strong state accountability systems have six essential elements (Reed, Scull, Slicker, & Winkler, 2012).

- Adoption of demanding, clear, and specific standards in all core content areas, and rigorous assessment of those standards;

- Reporting of accessible and actionable data to all stakeholders, including summative outcome data and other formative data to drive continuous improvement;
- Annual determinations and designations for each school and district that meaningfully differentiate their performance;
- A system of rewards and consequences to drive improvement at the school and district levels;
- A system of rewards and consequences to drive improvement at the individual student level; and
- A system of rewards and consequences to drive improvement at the individual teacher and administrator level.

Only a few educators have made efforts to examine the differences of the accountability systems and the effectiveness of the differences. Amrein and Berliner (2002) examined nine educational policies (high school graduation exams, high-stakes attached to tests, schools closed, principals replaced, grade-to-grade promotion, school choice, awards for schools, teachers, and students) of 27 states and calculated the number of policies that states implement. For example, Delaware, North Carolina, and Texas implemented six accountability policies, but Georgia, Minnesota, and Missouri executed only one policy. States that have high scores tend to have high school graduation exams and high-stakes tests. Researchers did not find consistent results that high-stakes tests and high school graduation exams increase students' performance.

Carnoy and Loeb (2002) created an index of accountability from 0 to 5, named the “strength” of the accountability system, using a database developed by the Consortium for Policy Research in Education (CPRE) which offers information on state testing and accountability



policies as of 1999-2000. “States receiving a zero do not test students statewide or do not set any statewide standards for schools or districts. ... States receiving a 5 test students in primary and middle grades, strongly sanction and reward schools or districts based on improvement in student test scores, and require a high school minimum competency exit test for graduation” (p. 311). For example, because Iowa and Nebraska did not have any state level accountability policies, their accountability strength was 0. However, Florida, New Jersey, North Carolina, and Texas implemented strong accountability policies, including high school exit exams, so they got 5. They found that accountability strength is significantly related to the mathematics accomplishment among eighth graders, especially for African American and Hispanic students, but are unrelated to students’ grade-to-grade progression rates. However, Carnoy and Loeb did not explain how they distinguished a 5 score from a 4 score.

Swanson and Stevenson (2002) examined the twenty-two states’ activities<sup>5</sup> related to standards-based assessment and accountability from studies conducted by the Council of Chief State School Officers and quantified the states’ activities, named an index of “policy activism”, using a Rasch measurement model. If states had performance standards in all academic subjects and statewide students’ performance assessments as of 1996, they gained high scores and are considered as high reform states. Maryland and Kentucky were the most active states in 50 states although Nebraska, Iowa, and Wyoming had low of standards activities. In the study of Swanson

---

<sup>5</sup> The twenty-two state policy activities were classified into four types: (1) content standards, (2) performance standards, (3) aligned assessments, and (4) professional standards. The activities are: (1) Math Document; (2) Science Document; (3) Math Standards; (4) Science Standards; (5) Language Arts Standards; (6) History Standards; (7) Math Innovativeness; (8) Science Innovativeness; (9) Recertification; (10) Licensure by Standards; (11) Certification Tests; (12) Major in Field; (13) Math Document; (14) Math Performance Levels; (15) Science Document; (16) Science Performance Levels; (17) Math Innovativeness; (18) Science Innovativeness; (19) Math Assessment; (20) Science Assessment; (21) Innovative Items; and (22) Innovative Tests.

and Stevenson, a state's policy activism does not influence standards-based instructional practices, such as emphasizing topic and skills, implementing pedagogical techniques, and employing classroom assessments. However, this study did not consider school-level variables as influential factors that reflected schools' organizational features.

Clarke, Pedulla, and colleagues created the Boston rating by using a three by three matrix of accountability: one dimension is the severity of accountability policies related to students, and the other dimension is the severity of accountability policies related to teachers, schools, and districts (M. Clarke et al., 2003; Pedulla et al., 2003). When states have regulated or legislated sanctions or decisions based on the states' test scores, the states are considered high stakes states. If states have promotion/retention or graduation policies, they are considered as high stakes states for students, and if states have accreditation, funds, or receivership, they are also considered as high stakes states for teachers, schools, and districts. Delaware, Florida, Georgia, and sixteen states implemented strong accountability policies not only for students, such as promotion/retention or graduation policies but also for teachers, such as accreditation or funds. However, Iowa had low policies for both and Idaho had low accountability policies for teachers and high accountability policies for students (Pedulla et al., 2003). Teachers in high-stakes states, compared to those in lower-stakes states, tend to feel more pressure, to use curriculum for aligning with the policies, to spend more time on instruction in testing areas, and to focus on test preparation.

Lee and Wong (2004) calculated the number of policies that states use based on three data set<sup>6</sup> and created a composite factor of state activism in accountability policy during the 1990s.

---

<sup>6</sup> Three data sets are (1) 1995-1996 data from the North Central Regional Education Laboratory (NCREL) and the Council of Chief State School Officers (CCSSO); (2) 1999 data from the Quality Counts (QC) report; and (3) 1999-2000 data from the Consortium for

Based on state activism, states were classified 50 states into three categories: states with strong accountability systems (12 states in the top quartile), those with moderate systems (25 states in the middle half), and those with weak systems (13 states in the bottom quartile). North Carolina, and Texas were states with strong accountability policies but Arkansas, Nebraska, and Wyoming were states with weak accountability systems. States with strong accountability systems tend to have assessment, report cards, performance rating of schools, rewards for successful schools, and reconstitution or major alteration of failing schools. However, many weak accountability states do not have direct incentives to schools in the form of performance ratings, rewards, assistance, and sanctions although they implement report cards for schools. Differences of accountability policies among states were not significantly related to the increase of mathematics and the reduction of racial and socioeconomic achievement gaps.

Nichols and his colleagues (2006) created ‘the accountability pressure rating (APR)’ based on an introduction essay, a reward/sanction sheet, and newspaper stories. When states feel high-stakes testing pressure, they gain high scores in APR. Texas had high-stakes testing pressure comparing to Wyoming. APR influences math NAEP performance only for certain subgroups, such as fourth-grade Hispanic and eighth-grade African American students, but it also increases the drop rate.

Finally, Wei (2008) generated the AMOs strength measured by the difference between starting points in 2003 and the intermediate goals in 2005. A larger difference means that it is more difficult for states to attain the goal, and that the states have stronger accountability systems. North Carolina and Missouri had high AMOs strength but Minnesota and New Mexico had low AMOs strength. States with strong AMOs strength tend to have a higher mathematics

achievement for fourth grade Hispanic students and eighth grade White students, but lower reading achievement for all eighth students and fourth grade white students.

Seven existing indexes of accountability are summarized by state in Table II-3. Most studies considered Maryland, Florida, and Texas as states with strong accountability systems but Iowa and Wyoming as start with weak accountability systems. However, there are differences between existing indexes. For example, Delaware and Michigan were considered as states with strong accountability systems in studies of Amrein and Clarke, but other studies did not.

Although six indexes exist, it is necessary to create and use a new accountability index for identifying states' accountability systems and understanding the effects of accountability on principals and teachers. In the early age of NCLB, few states implemented statewide tests and sanctions, so scholars consider the high-stakes tests and sanctions as indicators of strong accountability systems (Nichols et al., 2006; M. Clarke et al., 2003; Pedulla et al., 2003; Amrein & Berliner, 2002; Carnoy & Loeb, 2002). However, in 2012 most states have tests and sanctions. Therefore, statewide and sanctions cannot be longer suffice an indicator of high stakes states.

**Table II-3 Existing Indexes of Accountability**

State	Amrein& Berliner (2002)	Carnoy& Loeb (2002)	Swanson & Stevenson (2002)	Clarke et al (2003)		Lee & Wong (2004)	Nichols et al (2006)	Wei (2008)			
		Strength of accountability systems	Index of policy activism	Boston Rating		State activism in accountability policy	The accountability pressure rating	AMOs strength			
				Student	Teacher			G4M	G4R	G8M	G8R
AL	4	4	2.195	H	H	Strong	3.06	6.00	5.00	6.00	8.00
AK	-	1	-0.949	M	H	Weak	2.00	7.52	6.00	7.52	6.00
AZ	-	2	-0.395	H	M	Moderate	3.36	13.3	11.30	5.50	11.50
AR	-	1	-0.268	H	M	Weak	-	11.96	11.36	14.12	13.66
CA	5	4	0.090	H	H	Moderate	2.56	10.5	10.80	10.50	10.80
CO	5	1	0.662	L	H	Weak	-	5.13	5.64	9.83	6.35
CT	-	1	1.291	M	H	Moderate	1.60	9.00	11.00	9.00	11.00
DE	6	1	0.206	H	H	Weak	-	8.00	5.00	8.00	5.00
DC	-	-	-	-	-	Moderate	-	10.27	11.62	13.37	14.38
FL	5	5	-0.268	H	H	Strong	-	15.00	17.00	15.00	17.00
GA	1	2	0.662	H	H	Moderate	3.44	8.30	6.70	8.30	6.70
HI	-	1	0.320	L	M	Moderate	1.76	18.00	14.00	18.00	14.00
ID	-	1	-0.268	H	H	Weak	-	9.00	6.00	9.00	6.00
IL	-	2.5	0.320	M	H	Strong	-	7.82	6.64	7.82	6.64
IN	4	3	0.899	H	H	Strong	-	7.20	6.90	7.20	6.90
IA	-	0	-1.606	L	L	Weak	-	4.30	5.00	2.00	5.70
KS	-	1	0.320	L	H	Moderate	-	13.30	12.20	13.30	12.20
KY	4	4	1.969	L	H	Strong	0.54	7.73	5.25	8.35	5.44
LA	5	3	-0.026	H	H	Strong	3.72	11.7	10.50	11.70	10.50
ME	-	1	1.291	L	M	Weak	1.78	9.00	7.00	9.00	7.00
MD	5	4	2.459	H	H	Strong	2.82	12.20	14.00	16.80	13.70
MA	3	2	0.320	H	H	Weak	3.18	7.90	4.90	7.90	4.90
MI	5	1	0.434	M	H	Moderate	-	9.00	10.00	12.00	12.00
MN	1	2	-0.395	H	M	Moderate	-	3.50	3.00	3.50	3.00

Table II-3 (cont'd)

MS	2	3	0.547	H	H	Moderate	3.82	13.00	9.00	19.00	18.00
MO	1	1.5	1.023	L	H	Moderate	2.14	21.80	19.40	21.80	19.40
MT	-	1	-1.261	L	M	Weak	-	-	-	-	-
NE	-	0	-1.606	L	M	Weak	-	9.00	10.00	11.00	10.00
NV	4	1.5	0.320	H	H	Moderate	-	10.40	11.30	11.30	10.50
NH	-	1	1.153	L	M	Weak	-	9.00	10.00	9.00	10.00
NJ	3	5	-0.395	H	H	Strong	-	9.00	7.00	10.00	8.00
NM	5	4	0.779	H	H	Strong	3.28	4.21	3.77	4.21	3.77
NY	4	5	0.091	H	H	Strong	4.08	10.00	10.00	10.00	10.00
NC	6	5	1.597	H	H	Strong	4.14	6.40	7.80	6.40	7.80
ND	-	1	-0.026	L	M	Weak	-	13.60	8.70	16.70	9.70
OH	5	3	1.153	H	M	Moderate	-	10.00	10.00	10.00	10.00
OK	2	1	0.434	L	H	Moderate	-	10.00	10.00	10.00	10.00
OR	-	2.5	0.662	M	M	Moderate	-	10.00	10.00	10.00	10.00
PA	3	1	-0.661	M	H	Moderate	-	10.00	9.00	10.00	9.00
RI	-	1	0.091	L	H	Moderate	1.90	6.40	4.00	9.00	5.30
SC	5	3	0.899	H	H	Moderate	3.20	21.15	20.60	21.15	20.60
SD	-	1	-0.802	L	M	Moderate	-	9.00	6.00	9.00	6.00
TN	4	1.5	0.320	H	H	Moderate	3.50	7.00	6.00	7.00	6.00
TX	6	5	-0.661	H	H	Strong	4.78	8.60	6.20	8.60	6.20
UT	-	1	1.153	L	M	Moderate	2.80	7.00	6.00	7.00	6.00
VT	-	1	-0.268	L	H	Moderate	-	9.40	11.00	11.00	8.00
VA	2	2	0.547	H	H	Moderate	3.08	11.00	9.00	11.00	9.00
WA	-	1	0.206	H	M	Moderate	-	11.70	8.00	13.80	11.70
WV	3	3.5	0.899	M	H	Moderate	3.08	5.50	4.67	6.00	4.17
WI	-	2	-0.395	H	M	Moderate	-	10.50	6.50	10.50	6.50
WY	-	1	-0.949	L	M	Weak	1.00	12.70	11.60	12.45	10.92

Moreover, since 2002, states have continually modified their accountability policies for ten years. In 2002, there was no common core academic standard among 50 states but in 2012, forty-five states adopt them (Kober & Rentner, 2012). In addition, some states, including Indiana, North Carolina, and Oklahoma, have increased fourth and eighth grade reading and mathematics performance standards since 2002; however, Maine, South Carolina, and Wyoming's reading and mathematics performance standards have decreased (National Center for Education Statistics, 2011; Bandeira de Mello et al., 2009; McLaughlin et al., 2008; National Center for Education Statistics, 2007). The numbers of states that implement high school exit exams also have increased. In 2012, twenty-six states implement mandatory exit exams although in 2002 only nineteen states had (Zabala et al., 2008). Because the states' accountability policies in 2002 may be different from those in 2012, it may not good to examine the relationship between strength of states' accountability systems and principals' and teachers' responses suing existing indexes.

## **2) Major factors of states' accountability strength**

Based on these literatures about the difference of accountability system strength, I assume that accountability strength is determined by three major factors: the proficient performance standards, the annual measurable objectives (AMOs) strength, and the high school graduation exit exams.

### ***Proficiency performance standards***

NCLB requires each state to implement academic assessments and to review the annual progress of each school. Because assessment methods and standards were not specified in the law, each state decides on the high-stakes test that all students should take, and they set up its

content and performance standards accordingly.

Although content standards among 50 states have been similar since the introduction of the Kindergarten-12 Common Core State Standards in English and mathematics, there are huge variations in the proficiency performance standards that determine how well children are mastering the material in the state academic content standards. Some states, such as South Carolina and Massachusetts set up high proficiency performance standards, but other states, such as Tennessee and Oklahoma did not in 2003 (Peterson & Hess, 2005). Therefore, one student who passes the proficiency performance standards of Tennessee may not pass the standards of South Carolina.

The variations in the proficiency performance standards among states are remarkably shown by studies of the National Center for Education Statistics (NCES). NCES has studied the states' proficiency standards, considering the National Assessment of Educational Progress (NAEP) as a comparison metric<sup>7</sup> (National Center for Education Statistics, 2011; Bandeira de Mello et al., 2009; McLaughlin et al., 2008; National Center for Education Statistics, 2007). NCES found great variation in proficient performance standards in reading and mathematics across the states (see Appendix A and Appendix B). Massachusetts, Missouri, and South Carolina

---

<sup>7</sup> For a given subject and grade, the percentage of students reported in the state assessment to be meeting the standard in each NAEP school is matched to the point in the NAEP achievement scale corresponding to that percentage. "The method of obtaining *equipercentile equivalents* involves the following steps: (1) obtain for each school in the NAEP sample the proportion of students in that school who meet the state performance standard on the state's test; (2) estimate the state proportion of students who meet the standard on the state test, by weighting the proportions (from step 1) for the NAEP schools, using NAEP school weights; (3) estimate the weighted distribution of scores on the NAEP assessment for the state as a whole, based on the NAEP sample of schools and students within schools; and (4) Find the point on the NAEP scale at which the estimated proportion of students in the state who score above that point (using the distribution obtained in step 3) equals the proportion of students in the state who meet the state's own performance standard (obtained in step 2)"(Bandeira de Mello et al., 2009, p. 6).



had high performance but Mississippi, Oklahoma, and Tennessee had low performance standards in fourth and eighth reading and mathematics in the 2007-2008 school year (Bandeira de Mello et al., 2009).

Under the NCLB, all students should exceed proficiency performance standards by the 2013-2014 school year. Therefore, if states set up high proficient performance standards, the students in those states may have difficulties reaching these goals. For example, although two students acquire the same score of 185 in eighth grade mathematics, the student in Tennessee will pass the state's proficient performance but the other student in Massachusetts will not in 2009. Therefore, Massachusetts's principals and teachers will arguably focus more on students' test scores than Tennessee's principals and teachers do.

### ***Annual measurable objectives strength***

Under the NCLB, states explain how all students will meet proficient standards by the 2012 school year and will show the yearly annual measurable objectives (AMOs), which are the annual minimum required percentages of students who pass the states' proficient performance standards. As the first step, states set up a starting point<sup>8</sup>, which is an initial annual measurable objective for each subject area in 2002. It presents how many percentages of students meet or exceed the state's proficient standards that states established. However, New York, Oklahoma, and Vermont use their starting point and intermediate goals as scale score, not the percent proficient improvement. Based on states' workbooks, the starting point of eighth grade

---

<sup>8</sup> A starting point is based on "the higher of the percentage of students at the proficient level who are in (1) the state's lowest achieving group of students or (2) the school at the 20th percentile in the State, based on enrollment, among all schools ranked by the percentage of students at the proficient level" (NCLB, 2001 sec. 111 (b) (2) (E)).

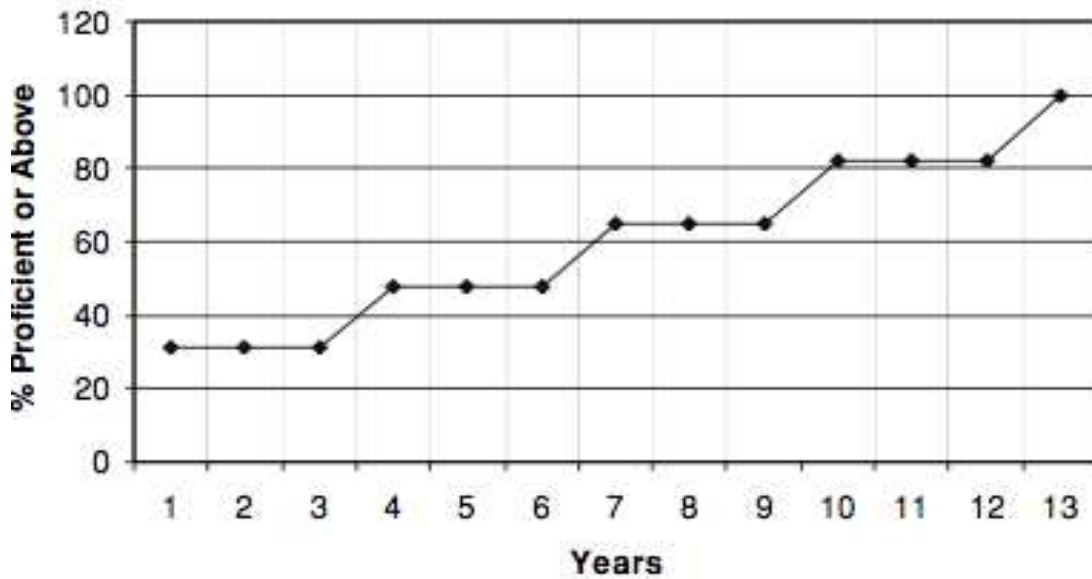
Mathematics in Arizona was 7%, while in Indiana it was 57.1%. Each state differently sets up its own starting points (see detail in Appendix C.).

Next, states set up the intermediate goals and illustrate how they will move from the starting point percent in 2002 to 100% in 2014. Intermediate goals, as prescribed by NCLB, must increase at least every third year, and each increase must be equal size. Because there is no specific regulation, each state chose intermediate goals every one, two, and three years, which leads to different trajectories: a straight-line pattern, a stair-step pattern (straight with plateaus), a front-loaded trajectory (larger increases for the early steps between plateaus), and a back-loaded trajectory (larger increases for the last steps between plateaus) (Porter, Linn, & Trimble, 2005)<sup>9</sup>. For example, as two figures show, Florida chose the stair-step approach with equal increases between steps and Michigan adopted the back-load trajectory approach.

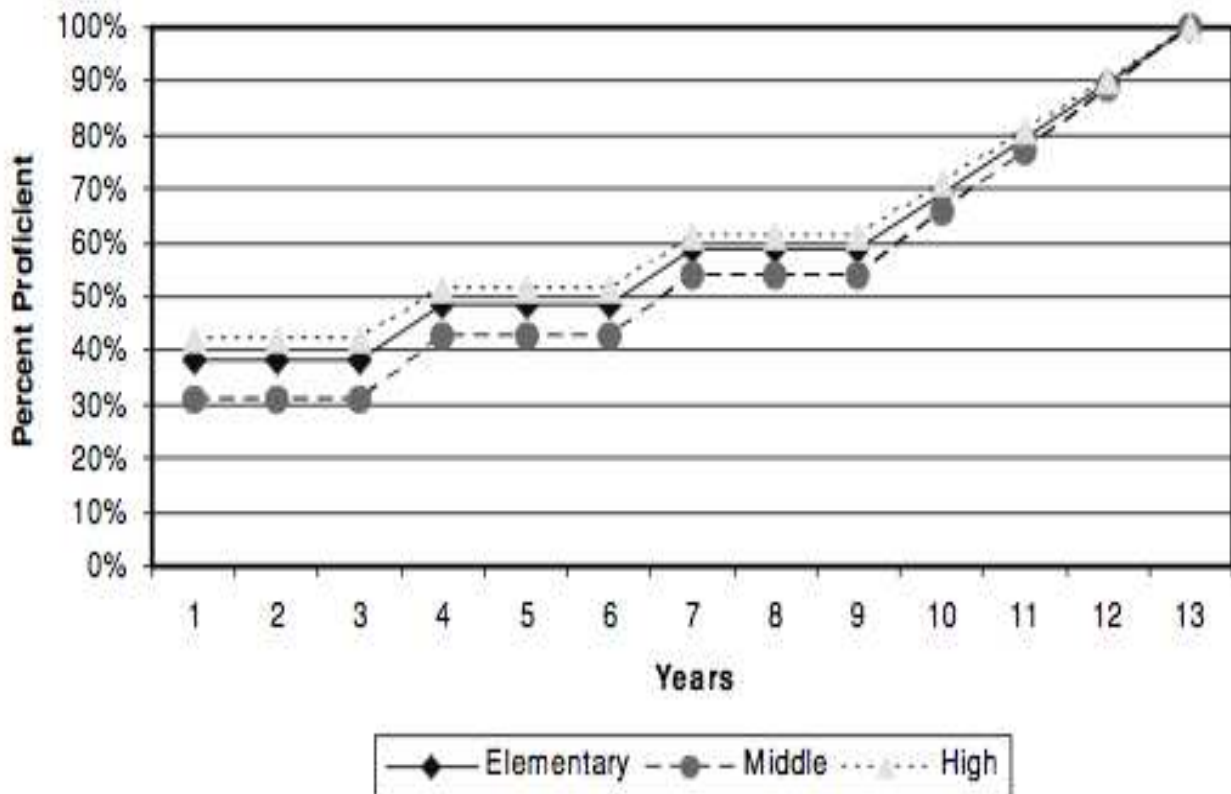
Appendix D shows 50 states' intermediate goals in 2007 or 2008. New Hampshire and Tennessee tended to set high intermediate goals, but California had low intermediate goals. Each state differently sets up its own intermediate goals.

---

<sup>9</sup> According to state's first workbook, most states selected a straight-line pattern and a stair-step pattern, but many states changed to back-loaded trajectory. In 2005, four states chose to use the straight-line pattern, nineteen states elected to use the stair-step pattern, and twenty-four states chose the back-loaded approach. No State chosen the front-loaded approach (Porter et al., 2005).



**Figure II-1 Florida Annual Measurable Objective for Reading**



**Figure II-2 Michigan Annual Measurable Objectives for Reading**

Citation from: Porter, A. C., Linn, R. L., & Trimble, C. S. (2005). The Effects of State Decisions About NCLB Adequate Yearly Progress Targets. *Educational Measurement, Issues and Practice*, 24(4), 36

Because each state establishes a different level of starting points and intermediate goals, some states have big differences between the starting point and the intermediate goal, but other states do not. If the differences are large, it may be difficult for students to reach the goals, and so principals and teachers focus on increasing students' test scores, spending more time and efforts on this. The states with large differences may have stronger accountability systems. However, if the differences are small, the states' accountability systems are not strong. For example, Florida had a big difference between the starting point and the intermediate goal, but Michigan had a small difference (see Figure II-1 and II-2). Therefore, in Florida students may be difficult to reach the goal and principals and teachers may more focus on students' test scores than Michigan's principals and teachers do.

### ***High school graduation exit exams***

Several states have implemented high school graduation exit exams, so students in the states need to take and pass the tests to receive a high school diploma. Generally, high school graduation exit exams are classified into two types: comprehensive exams and end-of-course exams (McIntosh, 2012). Comprehensive exams assess multiple subjects and are generally targeted at the 9th or 10th grade level, but end-of-course exams assess whether students master the content of specific high school classes (Zabala et al., 2007). In 2012, eight states used end-of-course exams and eighteen states implemented comprehensive exams as statewide and standardized final exams (McIntosh, 2012).

Because principals and teachers could get sanctions, if their students do not acquire grades high enough to pass the test, school staffs in states with mandatory tests for graduation may feel more pressure from accountability policies and spend more time on preparing the tests

and on teaching the curriculum related to the tests than do principals and teachers in states with no mandatory tests (Vogler, 2008; Bishop, Moriarty, & Mane, 2000).

### **3. Studies of Accountability Effects**

Since accountability systems implemented, a lot of educators have studied the effects of accountability systems in schools. This part will reveal the effects of accountability systems for students and teachers.

#### **1) The effects of accountability on students**

There are controversies about whether accountability policies increase or decrease students' accomplishment. Some found that accountability increases students' achievement (Jacob, 2005). Before NCLB, states with accountability policies tend to have higher students' academic achievement than states without the policies (Dee & Jacob, 2011). For example, prior to NCLB, Texas' various accountability policies, such as high stakes tests, encourage students to acquire higher NAEP test scores than other state's students (Grissmer, Flanagan, Kawata, & Williamson, 2000). After NCLB, students in states with high stakes testing acquire higher test scores than students in states without policies (Nichols et al., 2006; Hanushek & Raymond, 2005; Carnoy & Loeb, 2002). The researchers assume that strong accountability policies, such as high stake tests offer more pressure for increasing students' performance (Reback et al., 2011; Ladd & Lauen, 2010).

However, others claim that states with high stakes tests or high school graduation exit exams do not always have a high students' National Assessment of Educational Progress (NAEP), American College Test (ACT), Scholastic Aptitude Test (SAT), and Advanced Placement (AP) scores (S. S. Smith & Mickelson, 2000). Although Maryland implements stakes test, eighth<sup>h</sup> grade mathematics accomplishment did not increased (Amrein & Berliner, 2002). Although students in Texas receive the positive effects of strong accountability policies, their NAEP test

scores are not significantly higher than nationwide students' scores (Klein, Hamilton, McCaffrey, & Stecher, 2000). Moreover, there are few studies to find positive effects on reading achievement in the fourth and eighth grades (J. Lee & Reeves, 2012; Dee & Jacob, 2011; Schneider, 2011) and the number of schools that do not meet AYP has increased continuously since 2006 (Usher, 2012).

Researchers also have discussed the effects of accountability policies on an achievement gap. Some studies mention that accountability policies are effective in reducing inequalities in students' performance by race, socioeconomic status (SES), and achievement (Dee, Jacob, Hoxby, & Ladd, 2010; Henne & Jang, 2008; Hanushek & Raymond, 2004). The impact of accountability policies is larger for Black and Hispanic students than for White students: in fourth grade mathematics, Black and Hispanic students increased their NEAP test scores about 14.6 points and 9.8 points after implementation of NCLB although White students increased 4.9 points (Dee & Jacob, 2011). After NCLB, there is a reduction of an achievement gap between poor and non-poor students in fourth grade and eighth grade mathematics as well as between low performing and high performing students (J. Lee & Reeves, 2012; Ballou & Springer, 2009; Reback, 2008). Especially, reduction of achievement gap occurs in lower performing schools because the schools receive more pressure from accountability and pay more attention to minorities and economically disadvantaged students (Figlio, Rouse, & Schlosser, 2009; Springer, 2008).

However, other studies say that the accountability policies do not reduce students' academic achievement gaps because the effects of the policies differ by race and SES (Murnane & Levy, 2001; S. S. Smith & Mickelson, 2000). Because Black students receive more conventional teaching, such as lecture, recitation, and seat work than White students (Cox &

Witko, 2011; Diamond, 2007), they gain less effect from the policies than White students, and the achievement gap between students of different race is continued or expanded (J. Lee, 2006; Hanushek & Raymond, 2005; Hanushek & Raymond, 2004). Moreover, high school graduation test or requirements reduce low-achieving students' test scores (Dee, 2002; Jacob, 2001) and their higher order thinking skills (Rothstein, 2004). After NCLB, the achievement gap between low SES and high SES students has been changed significantly in both reading and math at fourth and eighth grades (J. Lee, 2006).

Additionally, studies produce different opinions about the unintended effects on students. Although critics argue that accountability policies decrease students' science or social studies achievements because teachers spend less time for non-test subjects (Cox & Witko, 2011; Diamond, 2007), supporters do not find any adverse impact on student non-test subject performance (Dee & Jacob, 2011; Winters, Trivitt, & Greene, 2010). Advocates observe that the strength of accountability does not produce significantly negative effects on graduation rates in high school (Carnoy, 2005), but opponents found that the policies lead more minority students to fail than White students (Haney, 2000). Strong accountability policies, including graduation exams and higher course requirements, enhance dropout rates (Jacob, 2001; Lillard & DeCicca, 2001) and reduce matriculation rates of low performance students (Bishop & Mane, 2001; Fuhrman et al., 1988).

According to studies focusing on students during the last ten years, recent studies by economists show that accountability is effective in increasing test scores, especially fourth grade mathematics test scores (Schneider, 2011), although earlier studies found negative effects of accountability. However, there is no agreement on achievement gap diminution (D. N. Harris & Herrington, 2006), nor on an English test score improvement (J. Lee & Reeves, 2012; Dee &



Jacob, 2011; Schneider, 2011). In addition, it is not clear whether the score increase comes from students' academic development or from the policy changes. Many states have modified their systems for more students to be counted as proficient (Rentner et al., 2006). They have lowered performance proficiency standards (J. Lee, 2010), enhanced minimum group size for analysis, used confidence intervals, permitted students to save test scores or retake tests, and modified many definitions, to all the state's advantage (Erpenbach, 2011; Forte & Erpenbach, 2006).

These studies excessively emphasize the effects of accountability policies on students' academic achievement and overlook the effects on students' non-academic outcomes, such as learning interest or attitude. Although NCLB highlights improvement in students' academic accomplishment, academic interest and learning attitude are more important and effective in improving students' academic outcomes in the long term (Hemmings & Kay, 2010). Excessive emphasis on achievement may be negative for students' cognitive development (Nichols & Berliner, 2007), and for students' widespread knowledge and skills acquisition (Koretz, 2005; Stecher, Chun, Barron, & Ross, 2000). Moreover, excessive emphasis on achievement leads students to avoid educational challenges and efforts (Dee, 2002), and weakens students' academic interest and inter-personal skills (Rothstein, Jacobsen, & Wilder, 2008).

## **2) The effects of accountability on teachers**

Researchers have studied the effects of accountability policies for teachers and some have found that accountability is effective to develop teaching quality. Under the rigorous accountability systems, teachers spend more time and effort on curriculum, teaching, and assessment (Kelley, 1999; Koretz & Training, 1996) because of local norms or agreements about accountability and professional practices (Swanson & Stevenson, 2002). Teachers reorganize their curriculum to fit accountability assessments (Srikantaiah, 2009; Hamilton et al., 2007;

Swanson & Stevenson, 2002; Kelley & Protsik, 1997), especially in states with strong accountability policies (Firestone, Mayrowetz, & Fairman, 1998). For example, teachers emphasize logical writing skills that are necessary for standards-based essays as well as high-order thinking skills, such as critical thinking and problem solving (Yeh, 2005; Wollman-Bonilla, 2004). In addition, teachers have modified their instructional methods and pedagogical techniques in order to align with the policies (Hamilton et al., 2008; Finnigan & Gross, 2007) and states' standards (Hamilton et al., 2007). Teachers employ various types of assessments (Swanson & Stevenson, 2002), emphasize classroom management (Koretz & Training, 1996), and apply data for decision-making and teaching practices (Srikantaiah, 2009; Hamilton et al., 2007; Kelley et al., 2000).

The accountability policies facilitate teacher collaboration and professional development. In the regular meetings required by the policies, teachers share knowledge of content and school reforms with colleagues, and thus establish a collaborative culture (Diamond, 2007; Kelley et al., 2000; Stecher et al., 2000). Also, these policies encourage teachers to participate in professional development and to create professional communities for improving their content knowledge and teaching skills (Srikantaiah, 2009; Libresco, 2005; Yeh, 2005; Firestone et al., 1998). Especially, when states implement state's achievement tests and the tests are aligned with state standards, teachers' participation in content-focused professional development is high (Phillips, Desimone, & Smith, 2011; Desimone, Smith, & Phillips, 2007).

However, other researchers found different results. Accountability policies emphasize only tested subjects and narrow instructional content (Anagnostopoulos, 2006; Booher-Jennings, 2005). Teachers increase their teaching time for test subjects (Cocke et al., 2011; Cox & Witko, 2011; Dee & Jacob, 2011; Reback et al., 2011; West, 2007), and decrease the time for non-tested

subjects (Kober, Chudowsky, & Chudowsky, 2008; Hamilton et al., 2007; Rouse, Hannaway, Goldhaber, & Figlio, 2007). Moreover, they limit the scope of class instruction to the specific content for testing, instead of general content in the subjects or higher-order skills (Srikantaiah, 2009; Hamilton et al., 2008; Diamond, 2007) and to “bubble kids” who are close to the proficient standard, instead of the other students (Neal & Schanzenbach, 2010; Reback, 2008; Hamilton et al., 2007; Booher-Jennings, 2005). These negative effects on teaching are shown more by teachers in states with strong accountability than by teachers in states with weak accountability (M. Clarke et al., 2003).

This content contraction leads to fragmented and teacher-centered instruction (Au, 2007). Due to the pressure of tests (L. M. McNeil, 2000; Sirotnik & Kimball, 1999), teaching styles are changed into “teaching to the test” (Diamond, 2007; Hoffman, Assaf, & Paris, 2001; Clotfelter & Ladd, 1996), which emphasizes memorization, recitation, and lecture (Diamond, 2012). Moreover, teachers are reluctant to implement innovative teaching practices (Hood, 2012; Martell, 2010; Crocco & Costigan, 2007) because teachers change their goals from improving students’ academic outcomes to reaching state level academic standards or receiving rewards (Finnigan & Gross, 2007; Booher-Jennings, 2005).

Moreover, accountability policies produce unintended negative effects. Teachers perceive low autonomy for important decisions in classrooms because of test grades and standards from states and districts (Diamond, 2012; Hood, 2012; Wills & Sandholtz, 2009; Garvin, 2007; Rouse et al., 2007). The policies tend to magnify teachers’ stress, frustration, and fatigue because of insufficient time to prepare for assessment (Finnigan & Gross, 2007; Abrams, Pedulla, & Madaus, 2003; Kelley et al., 2000), and because of conflicts between teachers’ own approaches and the enforced approaches to NCLB (Hamilton et al., 2007). The policies decrease the job

security of teachers (Opdycke, 2004), especially in low-performing schools (Reback et al., 2011), and they increase teacher turnover rates (Feng, Figlio, & Sass, 2010; Koretz & Training, 1996). High turnover rates can destroy collegiality and collaboration among teachers and produce isolated teachers (Rice & Malen, 2003).

Previous studies about teachers under accountability systems showed mixed effects (Hannaway & Hamilton, 2008; Hamilton et al., 2007). These studies have usually employed qualitative research methods that can reveal various teachers' activities. One study described both positive and negative effects of accountability policies. For example, although teachers align standards and instruction with performance-based tests, they also narrow curriculum and perform to test teaching.

Overall, accountability studies about teachers have overlooked teachers' sense-making processes (Schmidt & Datnow, 2005). These studies assume that teachers are passive, and thus they are affected by only accountability policies. However, teachers actively understand and arbitrate their policy environment and implementation based on their beliefs, knowledge, and prior experiences (Diamond, 2012; D. M. Harris, 2012; Rex & Nelson, 2004). Moreover, colleagues, principals, and school climate can help to produce a collective sense-making process (Louis, Febey, & Schroeder, 2005). Under the same accountability conditions, a principal's leadership can make a difference in teacher's motivation and teaching practice (Finnigan, 2012). Also, teachers' perceptions of and activities in response to of accountability policies are likely to be mediated by school organizations (Spillane et al., 2002).

#### **4. Principals' Response to Accountability**

As “street level workers” of state’s educational policies (Lipsky, 2010), principals can transform remote and intangible policies into closed and tangible outcomes (Rorrer & Skrla, 2005). Effective principals tend to have more influence on instruction and to support teachers’ learning to increase students’ academic achievement that is a major goal of accountability policies (Robinson, Lloyd, & Rowe, 2008; Bottoms, 2003).

##### **1) Having influence on instruction**

Each principal may have capacities and power to influence standards, curriculum, and instruction. Because principals understand the importance of standards to improve students’ performance (Printy, 2010), they can align their school’s standards with the state’s standards (Hamilton et al., 2007) and can establish performance standards using their students’ test scores (Lewis, 2010; Englert, Fries, Martin-Glenn, & Douglas, 2007; Ladd & Zelli, 2002; Spillane et al., 2002). To achieve the state’s performance standards, principals may match curriculum and instruction with state level standards or assessments (Hamilton et al., 2007; Marsh & LeFever, 2004). Moreover, principals can want to judge whether teachers are implementing teaching that can encourage students to meet the state’s standards. They may formally observe classrooms and evaluate teachers’ curriculum implementation (Gonzalez, 2012; Louis et al., 2010; Mojkowski, 2000).

Principals’ influence on instruction may be affected by their state. Some scholars claim that when states have more influence on developing standards, curriculum, and assessment, schools may be more accountable for student outcomes (Fuhrman & Elmore, 2004). However, others assert that as a state’s influence increases, principals’ and teachers’ influence may decrease

(Nance & Marks, 2008). Moreover, there can be differences among state's in their control of instruction, although most states have enacted legislation related to standards and curriculum (Louis et al., 2010). Principals in Massachusetts and Texas think that their states have more influence on instruction, but the principals in Nebraska and Montana do not think that about their state (Marks & Nance, 2007). Maryland provides principals with workshops and templates for standards, curriculum, and professional development (Jenkins & Pfeifer, 2012), so Maryland principals can have more power over instruction.

Under the situation that the accountability systems of states are not the same, I assume that there may be differences across the states in principals' influence on instruction. Although principals affect school standards, curriculum, and instruction, principals in states with strong accountability systems may feel more pressure from the accountability systems. This pressure may encourage principals to have more power for setting performance standards, defining curriculum, and evaluating teachers than principals in states with weak accountability systems.

## **2) Facilitating teacher learning**

Teacher learning is considered as an teachers' ongoing process of engagement in various activities that can produce their belief, knowledge, and instruction (Putnam & Borko, 1997). Usually, teacher learning can yield the change in knowledge and beliefs, the intentions for practice, the changes in actual teaching practices in a more permanent way, and the changes in emotions (Bakkenes, Vermunt, & Wubbels, 2010). The effects of teacher learning can lead successful students' academic outcomes (Lam, 2005).

To facilitate teachers' learning, principals can do two types of works. First, principals support teachers' professional work (Croft, Cogshall, Dolan, & Powers, 2010). Principals in

effective schools can reorganize faculty meetings to focus on professional development among teachers (Sanzo et al., 2011), including constructing common planning time for team meetings, securing additional time, and allocating school educational resources to support professional development (Graczewski, Knudson, & Holtzman, 2009; Kose, 2009; Arbogast, 2004; Youngs & King, 2002). Principals can also permit early dismissal of teachers to participate in professional development programs (Buchholz & List, 2009). Schools can provide substitute teachers so that staff can attend professional development programs offered by the district or state during the school day (Daniels, 2009; Roellke & Rice, 2008).

Second, principals can provide professional days before and during the school year. Lack of time has been cited as the most serious obstacle to the programs (Drage, 2010; Lind, 2007). Due to schools' schedules and teachers' classes, teachers can choose only a few professional development programs offered at different times and on different days, and it may be difficult for them to focus on professional development programs (Daniels, 2009). However, when principal provide professional days, teachers can obtain opportunities for professional development (Bubb & Earley, 2013).

In accountability contexts, principals' roles related to teachers' learning can be influenced by their state's systems (Spicer, 2008). States establish many regulations for teachers to participate in professional development, and they can provide financial funding for professional development (Boser, 2001; Dean, 2001). Just as each state has a different accountability system, the states' regulations of and supports for professional development may be not similar (M. Clarke et al., 2003). Kentucky has many requirements, 15 semester credit hours in the first five years (Loeb, Miller, & Strunk, 2009). Massachusetts, Kentucky, and North Carolina, which have high-stakes tests, offer more financial resources than Kansas, which has low-stakes tests (B.

Berry et al., 2003; M. Clarke et al., 2003).

Given that states' accountability systems are not the same, I assume that there may be differences in principals' facilitating teachers' learning. Principals in states with strong accountability systems may feel more pressure from the accountability systems and provide more supportive strategies than principals in states with low accountability systems in order to facilitate teachers' learning and to encourage teachers to participate in professional development programs.



## **5. Teachers' Responses in Accountability**

In the prior part, I explained principals' responses to accountability, including emphasizing standards, curriculum, and instruction, and facilitating teachers' learning. In this section, I will describe teachers' responses, teacher autonomy and their professional development participation time, which can be influenced by principals and states (Murnane & Papay, 2010).

### **1) Teacher autonomy**

In American education, teacher autonomy is considered as an important influential factor for school education, although it can produce a un-collaborative school climate that encourages teachers to work alone (O'Hara, 2006). First, teacher autonomy, as one source of intrinsic motivation, can improve teachers' professionalism. When teachers have opportunities to participate in policies, such as textbook and curriculum adoption, they can consider themselves as a major person (Kelley & Protsik, 1997) and they can consider teaching as interesting and meaningful professional work (Roth, Assor, Kanat-Maymon, & Kaplan, 2007; Pearson & Moomaw, 2005). Second, teacher autonomy can decrease stress and increase job satisfaction (Pearson & Moomaw, 2005). Teachers who are more autonomous in their classrooms may have high satisfaction and remain in their teaching jobs (Pearson & Moomaw, 2006; Rudolph, 2006; Brunetti, 2001).

When we study teacher autonomy, principals can be considered as essential, because principals may influence teachers' instruction (Printy, 2010). Teachers' autonomy can be greater or less based on how principals handle external requirements and expectation (Rudolph, 2006). When principals implement 'tight and direct control,' teacher autonomy may be diminished (Eden, 2001). For example, when principals choose the curriculum and mandate instruction,

teachers can be limited in using their own favorite curriculum and new instructional methods. However, principals can also encourage teacher autonomy (Pearson, 1995). When principals give more opportunities for participation in major decisions, understand teachers' conditions and needs, establish a school climate which supports teacher autonomy, and assign autonomy to teachers, teachers can gain improved autonomy (Assor & Oplatka, 2003). Such teachers can feel that their principal protects them from the pressure of their state's administration and produces a school environment in which teachers can implement their autonomy (Byrne, 2009; Crocco & Costigan, 2007).

It is assumed that educational reforms threaten teacher autonomy (Quiocho & Stall, 2008; Spillane et al., 2002; Brunetti, 2001). Under the accountability policies, standards, contends, and curriculum for classroom learning are given and teachers had little flexibility in the content they taught (Desimone, 2013). For example, many educational works, such as curriculum, texts, class size, scheduling, and space allocations, may be controlled by legislatures rather than by teachers (Pearson & Hall, 1993). To reach a state's performance standards, teachers may abandon their curriculum and the teaching practices that are best for their students, and they may diminish their creativity, choice, and spontaneity (Hood, 2012; Martell, 2010; Wills & Sandholtz, 2009; Garvin, 2007).

Other opinions also exist. In schools that are loosely coupled systems, accountability can produce recoupling between policies and classrooms (Hallett, 2010). Because of the process of creating tight couplings, teachers may follow accountability regulations in some areas: however, in other areas where loose couplings still exist, teachers can maintain their autonomy. For example, teachers can change their curriculum based on the state's content standards, but they may continue their instruction with autonomy (Diamond, 2012; Spillane, Parise, & Sherer, 2011).

Teachers may need and seek to find a balance point between accountability and professionalism.

Under the situation that the states' accountability systems are not the same, I assume that there might be differences in principals' influence on standards, curriculum, and instruction, and this difference in influence can affect teacher autonomy. These differences in principals' activities can be related to teacher autonomy.

## **2) Teachers' professional development**

In the era of accountability, the importance of professional development has increased because professional development can be an effective method to improve teachers' content knowledge, their teaching capacities, and high order thinking skills that are related to students' performance. First, from professional development teachers can acquire content knowledge (Youngs & King, 2002; Ball & Cohen, 1999) and problem-solving abilities (Jasper & Taube, 2004). Second, professional development can improve teachers' instruction (Hill, 2007; Lambert, 2003; Garet, Porter, Desimone, Birman, & Yoon, 2001) and develops alternative student assessments in their classrooms (Sato, Wei, & Darling-Hammond, 2008; Desimone, Porter, Garet, Yoon, & Birman, 2002). Third, schools where teachers gain opportunities to participate in professional development about students' performance and educational policies can have higher student educational outcomes than schools where teachers do not have these opportunities (Louis et al., 2010; Yoon et al., 2008; Joyce & Showers, 2002).

Principals have been considered as essential beings for teachers' professional development (Youngs & King, 2002; Hallinger & Murphy, 1986). First, as a builder and designer of professional development, principals can design professional development programs based on school visions, and they evaluate the programs (Kose, 2009; Lindstrom & Speck, 2004).

Moreover, principals in effective schools can reorganize faculty meetings to focus on professional development among teachers (Sanzo et al., 2011). When principals establish more consistent visions, teachers can coherently participate in professional development programs (Graczewski et al., 2009).

Second, principals can create school contexts that encourage teachers to actively participate in professional development programs (Wahlstrom & York-Barr, 2011; Rice & Malen, 2003). Principals can secure additional time, find additional funds for workshops and conferences, and allocate school educational resources to support professional development (Graczewski et al., 2009; Kose, 2009; Arbogast, 2004; Youngs & King, 2002). They can offer opportunities for teachers to connect with various organizations from outside, such as local universities and nonprofit organizations for external assistance (Sebring & Bryk, 2000), and to participate in decision-making processes related to professional development (Newmann, King, & Youngs, 2000).

Teachers' participation in professional development can be influenced by their state's educational policies. A state's strong tasks accountability atmosphere can lead teachers to participate in professional development programs (Desimone et al., 2007), so when states use criterion-referenced assessments that are aligned to state standards in mathematics at the high school level, teachers' participation time in content-focused professional development can increase (Phillips et al., 2011; Desimone et al., 2007). Also, 24 states provide some money for professional development, and 38 states make regulations that teacher should participate in professional development to maintain their licenses (Boser, 2001).

Under the situation that the states' accountability systems are not the same, I assume that there might be differences in principals' facilitating teachers' learning. Differences in principals'

activities can influence differences of teachers' participation time in professional development.

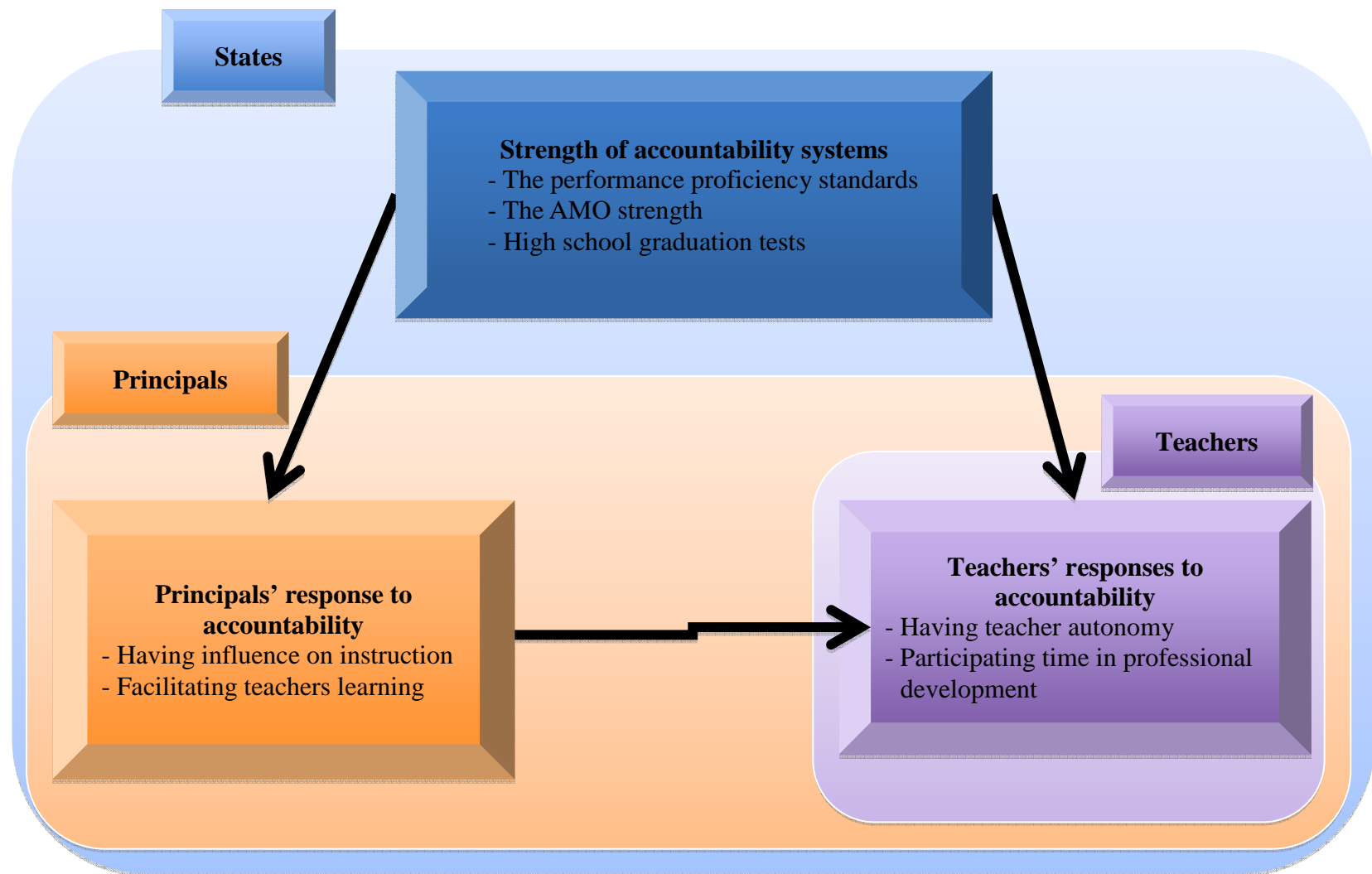
## **CHAPTER THREE**

### **III. METHODOLOGY**

In this methodology chapter, I will explain a conceptual model, research questions, hypotheses, data sets, variables, and analysis methods.

#### **1. Conceptual Model**

Based on literature review, I established a simple conceptual model like Figure III-1 comprised of three parts: states, principals, and teachers. I found that each state has different accountability policies: the proficiency performance standards, the annual measurable objectives (AMO) strength, and high school graduation exit exams . Based on these findings, I assumed that these dissimilar policies could make the different level of accountability strength. This states' accountability strength can affect principals' responses: principals' influence on instruction and their facilitation of teachers' learning. This states' accountability strength can also influence teacher autonomy and their participation time in professional development programs. Principals' perception and activities may affect teachers' educational activities, such as instruction, selection of curriculum, and participation in professional development.



**Figure III-1 A Conceptual Model**

**For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation**

## 2. Research Questions and Hypotheses

Based on the conceptual model, I created the two research questions and nine hypotheses.

**The first research question:** What is the relationship between strength of states' accountability systems and principals' responses: their influence on instruction and their facilitation of teacher learning?

- a) Are there differences in principals' responses to accountability strength of 51 states?
- b) Which states' accountability strength factors affect principals' responses?
- c) Which principals' individual factors and school environmental factors affect principals' responses?

**Hypothesis 1:** States' high proficient performance standards, AMO strength, and high school graduation exit exams will be negatively and significantly correlated with principals' influence on instruction.

**Hypothesis 2:** States' high proficient performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with principals' support of professional work.

**Hypothesis 3:** States' high proficient performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with principals' provision of professional days before the school year.

**Hypothesis 4:** States' high proficient performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with principals' provision of professional days during the school year.



**The second research question:** What is the relationship between strength of states' accountability systems and teachers' responses: teacher autonomy and their participation in professional development programs?

- a) Are there differences in teachers' responses to accountability strength of 51 states?
- b) Which states' accountability strength factors affect teachers' responses?
- c) Which teachers' individual factors, principals' individual factors, and school environmental factors affect teachers' responses?

**Hypothesis 5:** States' high proficient performance standards, AMO strength, and high school graduation exit exams will be negatively and significantly correlated with teacher curriculum autonomy.

**Hypothesis 6:** States' high proficient performance standards, AMO strength, and high school graduation exit exams will be negatively and significantly correlated with teacher instructional autonomy.

**Hypothesis 7:** States' high proficient performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with teachers' participation time in professional development programs related to content.

**Hypothesis 8:** States' high proficient performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with teachers' participation time in professional development programs related to instruction.

**Hypothesis 9:** States' high proficient performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with teachers' participation time in professional development programs related to classroom management.

### **3. Data**

To respond to these research questions, information about states, district, schools, and principals was needed. First, state level information related to accountability systems came from states' Consolidated Application Accountability Workbook. The workbook is organized using ten principles of accountability and explains a plan how each state implements a statewide accountability system that included all public schools and all students in the schools. Their workbooks illustrate proficiency performance standards, starting points, intermediate goals, and assessment systems.

Second, School principals and school information came from the National Center for Educational Statistic's (NCES) School And Staff Survey (SASS) 2007-2008. SASS, as a set of questionnaires of teachers, principals, schools, and districts, provides descriptive data in the context of elementary and secondary education. SASS includes teacher education, certification, school climate, school size, and student population in 50 states.

SASS has four components from the school questionnaire, the teacher questionnaire, the principal questionnaire, and the school district questionnaire by sent to respondents in public, private, and Bureau of Indian Education/tribal schools. I handled only public school, principals, and teachers because public schools may be more influenced by state accountability systems than those private schools (McDonald, 2002) because public schools should follow mandates, rules, and regulations which education agencies set in order to maintain educational funding (Rudolph, 2006).

Since 1987, SASS has been investigated: 1987-1988 school year, 1990-1991 school year, 1993-1994 school year, 1999-2000 school year, 2003-2004 school year, 2007-2008 school year, and 2011-2012 school year. The data set in the 2003-2004 may not represent states'

accountability conditions under NCLB begun in 2002. Also, because since 2011 states have gained flexibility of accountability, data set in the 2011-2012 may not present states' accountability systems. Based on this reason, among the seven data set, I used a data set in the 2007-2008 SASS that include 9,800 public schools, 9,800 public school principals, and 47,600 public school teachers.

In this study, I gave attention to the responses of principals and teachers in the secondary schools. Under the accountability policies, secondary schools may receive more concern than elementary schools. Secondary schools tend to be a large complex organization because of specialized content focus, so teachers in these schools may be likely to have more professional autonomy (Gross & Goertz, 2005). Moreover, high school graduation exit exams can influence principals, teachers, and students in secondary schools. The public secondary school data include 2,847 principals and 19,973 teachers.

In the public secondary school data, there are 2,112 male principals and 735 female principals. The SASS public secondary school data has more white principals (2,526) than non-white principals (321). The number of suburban schools principals (1,316) is more than the number of rural and urban schools principals. The characteristics of principal data set are shown in Table III-1.

**Table III-1 The Characteristics of Principal Data Set**

Gender		Race		School region			Total
Male	Female	Non-white	White	Urban	Suburban	Rural	
2,112	735	321	2,526	580	1,316	951	2,847

In the public secondary school teacher data, there are 8,350 male and 11,623 female

teachers. There are more white principals than non-white teachers. The number of suburban schools teachers is more than the number of rural and urban schools teachers. The characteristics of principal data set are shown in Table III-2.

**Table III-2 The Characteristics of Teacher Data Set**

Gender		Race		School region			Total
Male	Female	Non-white	White	Urban	Suburban	Rural	
8,350	11,623	1,701	18,272	580	1,316	951	19,973

According to the secondary schools data, among 19,973 teachers, the number of teachers who responded that they teach eighth grade is 3,272. About four hundred fifty teachers said that they teach English field, such as English, reading, and speech. Another about hundred fifty teachers teach mathematics such as algebra, calculus, and geometry. About two thousand three hundred fifty teachers mentioned that they teach eighth grade, not English or mathematics. The Table III-3 shows the distribution of teachers by subjects and grades.

**Table III-3 The Characteristics of Teacher Data Set**

Grade	Test subjects		Non-test Subjects	Total
	English	Mathematics		
Seventh grade	447	371	2,050	2,868
Eighth grade	452	471	2,349	3,272
Ninth grade	1,742	1,987	9,222	12,951
Tenth grade	1,915	2,219	10,605	14,739
Eleventh grade	2,011	2,214	10,887	15,112
Twelfth grade	1,941	2,085	10,588	14,614

## **4. Variables**

### **1) The strength of states' accountability systems**

Accountability strength was comprised of three factors: proficiency performance standards, AMO strength, and high school graduation exit tests. The first factor was the level of proficiency performance standards, which is the corresponding NAEP score based on states' proficiency performance standards. There are two types of proficiency performance standards: reading and mathematics. Original data of proficiency performance standards for math and reading were non-normal distribution, which cannot produce reliable results. I changed these two original data into standardization data and gained two types of z-scores; proficiency performance standards z-score for reading and proficiency performance standards z-score for mathematics. After acquirement of these two z-scores, I acquired mean of two z-scores of proficiency performance standards for reading and mathematics. The higher the scores, the more difficult states reach the goals and the stronger states have accountability systems. I obtained this information from Bandeira de Mello's (2009) reports. However, assessment data of Nebraska and Utah State were not available.

The second factor was AMO strength, which is the mean difference of starting points for reading and mathematics in 2002 and intermediate goals in 2007 for reading or mathematics assessment in eighth grade. However, New York, Oklahoma and Vermont used their starting point and intermediate goals as scale score, not the percent proficient improvement. I divided the difference between starting point and intermediate goals on the scale score by the maximum scale score on the test to calculate the accountability strength. Because like proficiency performance standards, original data of AMO strengths were abnormal distribution, I gained the mean of two standardization data of AMO strength z-score for reading and AMO strength z-score

for mathematics. The larger differences of between starting points and intermediate goals may be more difficult for students to approach at goals and the stronger states have accountability.

Starting points and intermediate goals were shown in section 3 of state's accountability workbooks.

**Table III-4 The Strength of States' Accountability Systems**

	Proficiency performance standards	AMO strength	High school graduation exit exams		Proficiency performance standards	AMO strength	High school graduation exit exams
AL	-1.00	-0.66	1	MT	0.51	-0.57	0
AR	-0.63	-0.62	1	NE	-	0.23	0
AZ	-0.11	0.89	1	NV	-0.07	0.02	1
AK	0.34	-0.14	0	NH	0.83	-1.38	0
CA	-	0.32	1	NJ	0.28	0.22	1
CO	-0.94	-0.18	0	NM	0.57	0.27	1
CT	-0.65	0.15	0	NY	0.59	-1.89	1
DE	-0.15	-0.46	0	NC	-1.03	-0.91	1
FL	0.43	0.77	1	ND	0.48	0.79	0
GA	-2.02	-0.69	1	OH	-0.39	-0.30	1
HI	0.77	1.41	0	OK	-1.21	-1.88	0
ID	-0.63	-0.31	1	OR	-0.10	0.19	0
IL	-1.00	0.45	0	PA	-0.01	0.07	0
IN	0.04	-0.43	1	RI	0.55	-0.44	0
IA	0.01	-0.48	0	SC	2.66	2.45	1
KS	-0.18	0.75	0	SD	0.14	0.36	0
KY	0.48	-0.13	0	TN	-2.46	-0.89	1
LA	-0.11	0.41	1	TX	-0.92	-0.36	1
ME	1.07	0.24	0	UT	-	-0.55	0
MD	0.41	1.52	0	VT	1.07	-1.89	0
MA	1.29	0.06	1	VA	-0.62	-0.18	1
MI	-0.63	0.50	0	WA	0.79	2.06	1
MN	1.21	-1.32	1	WV	-1.18	-1.37	0
MS	-0.10	1.85	1	WI	-0.81	-0.16	0
MO	1.56	1.70	0	WY	0.34	0.52	0

The last factor was high school graduation exit exams. Although many states did not have high school graduation exit exams, some states implemented high school graduation exit exams, such as comprehensive exams and end-of-course exams. Therefore, I coded states with high school graduation exit exams as 1 and states without exams as 0. I obtained this information from Zabala (2007)'s report "State High School Exit Exams: Working to Raise Test Scores". Accountability strength of each stat was shown in Table III-4.

## **2) Principals' responses**

Principals' responses include four factors: principals' influence on instruction, their support of professional work, their provision of professional days before the school day, and their provision of professional days during the school day. These variables come from principal questionnaire in SASS 2007 data set.

Although instruction is a teachers' work, principals can also influence teachers' instruction in three fields: setting performance standards for students of this school, establishing curriculum at this school, and evaluating teacher's instruction at this school. These questions were measured by a 5-point of Likert-type scale.

Principals encourage teachers to participate in professional development through facilitating teacher learning and provision of professional days. Facilitating teacher learning are reducing teachers' work, employing substitute teachers to cover teachers' classes, using common planning time, and operating early dismissal or late start for students.

Provision of professional days happens in the beginning of the students' school years and during the students' school years. These aspects were measured by whether or not schools use (yes = 1 / no = 0).

### **3) Teachers' responses**

Teachers' responses were comprised of teacher autonomy and teachers' participation time in professional development programs. Teachers usually can have two types of autonomy: teacher curriculum autonomy and teacher instructional autonomy. Teacher curriculum autonomy is selecting textbooks and other instructional materials and selecting content, topics, and skills to be taught. Teacher instructional autonomy is selecting teaching techniques, evaluating and grading students, disciplining students, and determining the amount of homework to be assigned. These questions were measured by a 5-point of Likert-type scale.

Teachers can participate in various professional developments including the content of the subjects they teach; reading instruction; student discipline and management in the classroom. Although I could make a model that has one dependent variable as an average of three fields of professional development participation time, I can assume that the taking of one kind of professional development may be unrelated to the taking of other kinds of professional development programs (Desimone et al., 2007). Therefore, I analyzed three kinds of professional development participation time as dependent variables. Teachers answered their participation time in six professional development programs in the past 12 months. These questions were measured by a 4-point scale a) 8 hours or less, b) 9-16 hours, c) 17-32 hours, and d) 33 hours or more). Major variables are shown in Table III-5. These variables come from teacher questionnaire in SASS 2007 data set.



**Table III-5 Variables of Principals' and Teachers' responses**

<b>Variables</b>		<b>Questions</b>
Principals' influence on instruction		<p>How much actual influence do you think each group or person has on decisions concerning the following activities?</p> <ul style="list-style-type: none"> <li>i. Setting performance standards for students of this school</li> <li>ii. Establishing curriculum at this school</li> <li>iii. Evaluating teachers of this school</li> </ul>
Principals' facilitation of teacher learning	Support of professional work	<p>Are the following used to provide teachers in this school with time for professional development during regular contract hours?</p> <ul style="list-style-type: none"> <li>i. Substitute teachers to cover teachers' classes</li> <li>ii. Early dismissal or late start for student</li> <li>iii. Common planning time for teachers for professional development</li> <li>iv. Reduced teacher work loads</li> </ul>
	Provision of professional days	<p>Are the following used to provide teachers in this school with time for professional development during regular contract hours?</p> <ul style="list-style-type: none"> <li>i. Professional days built in before the beginning of the students' school year</li> <li>ii. Professional days built in during the school year</li> </ul>
Teacher autonomy	Curriculum	<p>How much actual control do you have in our classroom at this school over the following areas of your planning and teaching?</p> <ul style="list-style-type: none"> <li>i. Selecting textbooks and other instructional materials</li> <li>ii. Selecting content, topics, and skills to be taught</li> </ul>
	Instruction	<p>How much actual control do you have in our classroom at this school over the following areas of your planning and teaching</p> <ul style="list-style-type: none"> <li>i. Selecting teaching techniques</li> <li>ii. Evaluating and grading students</li> <li>iii. Disciplining students</li> <li>iv. Determining the amount of homework to be assigned</li> </ul>
Teachers participation time in professional development		<p>In the past 3years, how many hours did you spend on these activities?</p> <ul style="list-style-type: none"> <li>i. Content of the subjects you teach</li> <li>ii. Reading instruction</li> <li>iii. Students discipline and management in the classroom</li> </ul>

#### **4) Control variables**

To respond four research questions, I used control variables in the level of principals, schools, and teachers. Principals' control variables were gender, race, educational background, the years as principals, ASPIRING participation, and previous participation in professional development. High scores of educational background mean that the principals may have high degree. The years as principals are the years principals serve as the principal of a current school and any other schools. Principals' control variables come from principal questionnaire in SASS 2007 data set.

Schools' control variables were region, size, and social economic status (SES). School region was classified into large or mid-size central city, urban fringe, and small town or rural area. I let the "Large or mid-size central city" category be the reference category, and create two dummy variables: urban fringe and small town or rural area. School size was measured by the number of student who is enrolled in the schools. School SES was inversely measured by the number of students who participate in the federal free or reduced-price lunch programs. High scores of school SES means that the schools have few students who participate in free or reduced-price lunch programs. These variables come from school questionnaire in SASS 2007 data set.

There were three types of school climate: teachers' shared responsibility, student learning attitude, and schools' resource adequacy. Three climates were measured four, two, and four teachers' questions that were measured by a 5-point of Likert-type scale. When the scores of teachers' shared responsibility are high, teachers perceived that they have high-shared responsibility about accountability systems. High scores of student learning attitude mean good student learning attitude. Low schools' resource adequacy scores means that school have less

hygiene factors which affect dissatisfaction although these factors do not motivate teachers.

Teachers' control variables were gender, race, educational background, teaching years, high-qualified teachers, and eighth grade test subject teacher. Eighth grade test subject teachers were eighth teachers who teach Mathematics and English. These variables come from teacher questionnaire in SASS 2007 data set. Specific questions were shown in Table III-6.

**Table III-6 Control Variables**

<b>Variables</b>	<b>Questions</b>
<b>Principals / Schools</b>	
Gender	Male: 0 / Female: 1
Race	Non-white: 0 / White: 1
Educational background	Below Master: 1 / Specialist: 2 / Doctoral: 3
The years as principal	The years as principals serve as the principal of a current school and any other schools
ASPIRING program	No: 0 / Yes: 1
Professional development participation	No: 0 / Yes: 1
Suburban	Large or mid size central city: 0 / Urban fringe: 1
Rural	Large or mid size central city: 0 / Small town or rural area: 1
Size	The number of students who enrolled in the school
School SES	The number of students who participate in the federal free or reduced-price lunch program (Inversely coding)
Teachers' shared responsibility	<ul style="list-style-type: none"> <li>i. Rules for student behavior are consistently enforced by teachers in this school, even for students who are not in their classes.</li> <li>ii. Most of my colleagues share my beliefs and values about what the central mission of the school should be.</li> <li>iii. There is a great deal of cooperative effort among the staff members.</li> <li>iv. In this school, staff members are recognized for a job well done.</li> </ul>
Student learning attitude	<ul style="list-style-type: none"> <li>i. The level of student misbehavior in this school (such as noise, horseplay or fighting in the halls, cafeteria, or student lounge) interferes with my teaching.</li> <li>ii. The amount of student tardiness and class cutting in this school interferes with my teaching.</li> </ul>
Schools' resource adequacy	<ul style="list-style-type: none"> <li>i. I am satisfied with my teaching salary.</li> <li>ii. Necessary materials such as textbooks, supplies, and copy machines are available as needed by the staff.</li> <li>iii. Routine duties and paperwork interfere with my job of teaching.</li> <li>iv. I am given the support I need to teach students with special needs.</li> </ul>
<b>Teachers</b>	
Gender	Male: 0 / Female: 1
Race	Non-white: 0 / White: 1
Educational background	Bachelor: 0 / Master: 1
Teaching years	The years as teachers serve as the teachers
High qualified teachers	No: 0 / Yes: 1
Eighth grade & test subject	The other grade: 0 / Eighth grade & English or Math: 1

## 5. Analysis of Principal's Responses to Accountability Policies

In order to examine whether or not there is relations between the strength of states' accountability systems and principals behaviors related to professional development and instructions, I used 2-level hierarchical linear modeling (HLM). Principals are nested within their states. HLM can reveal these hierarchical features and enable researchers to examine relationships involving predictors at two or more levels (Davison, Kwak, Seo, & Choi, 2002; Whitener, 2001).

**Table III-7 Descriptive Statistics for the 2-level Analysis Variables**

	Variable name	N	Mean	SD	Min	Max
Level-1	Influence on instruction	2,640	3.65	0.42	1.00	4.00
	Support of professional work	2,557	2.37	0.99	0.00	4.00
	Provision of professional days before school year	2,557	0.96	0.18	0.00	1.00
	Provision of professional days during school year	2,557	0.92	0.26	0.00	1.00
	Principal gender	2,640	0.26	0.44	0.00	1.00
	Principals race	2,640	0.88	0.32	0.00	1.00
	Principals educational background	2,640	1.49	0.67	1.00	3.00
	The years as principal	2,640	8.44	6.76	1.00	45.00
	ASPIRING program	2,640	0.52	0.50	0.00	1.00
	Professional development participation	2,640	0.98	0.15	0.00	1.00
	Suburban	2,640	0.46	0.50	0.00	1.00
	Rural	2,640	0.34	0.47	0.00	1.00
	School size	2,513	3.63	1.47	1.00	5.00
	School SES	2,505	2.50	1.52	0.00	5.00
	Teachers' shared responsibility	2,524	2.94	0.37	1.25	4.00
	Student learning attitude	2,524	2.71	0.52	1.00	4.00
	Schools' resource adequacy	2,524	2.65	0.32	1.25	4.00
Level-2	The proficiency performance standards	47	-0.01	0.93	-2.46	2.66
	AMO strength	47	0.00	0.99	-1.89	2.45
	High school graduation exit exam	47	0.47	0.50	0.00	1.00

Among the 50 states, California, Nebraska, and Utah State did not have proficiency performance standards, so three states were excluded from level 2 HLM analysis. About 2,600 principals in 47 states were analyzed in 2-level analysis. Descriptive statistics for the 2-level analysis variables appear in Table III-7.

First of all, I analyzed a fully unconditional model - one-way ANOVA with random effects – with principals’ influence of instruction and their support professional development to estimate the proportion of within- and between-group variability in the dependent variable (Raudenbush & Bryk, 1992). A fully unconditional model represented below.

Principal level: Influence on instruction<sub>ij</sub> = B0 + R

Support of professional work<sub>ij</sub> = B0 + R

Provision of professional days before the school year<sub>ij</sub> = B0 + R

Provision of professional days during the school year<sub>ij</sub> = B0 + R

State level: B0 = G00 + U0

Influence on instruction<sub>ij</sub> = The level of influence on instruction of principal i in state j

Support of professional work<sub>ij</sub> = The level of support of professional work of principal i  
in state j

Provision of professional days before the school year<sub>ij</sub> = The level of provision of  
professional days before the school year of principal i in state j

Provision of professional days during the school year<sub>ij</sub> = The level of provision of  
professional days during the school year of principal i in state j

B0 = Each state's mean of principals influence on instruction, facilitating teacher learning, or provision of professional days

G00 = Grand mean of principals influence on instruction, facilitating teacher learning, or provision of professional days

R = The principal level variance

U = The state level variance

This fully unconditional model analysis can yield an intra-class correlation coefficient (ICC), which is “the proportion of the variance in the outcome variable that is between the second-level units” (Kreft & Leeuw, 1998, p. 9). In this study, ICC represented the proportion of variance in principals' responses between states. The formula for ICC is  $ICC = \hat{\tau}_{00} / (\hat{\tau}_{00} + \hat{\sigma}^2)$ , where  $\tau_{00}$  is the variability of  $Y_{ij}$  at the first level, and  $\sigma^2$  is the variance of  $u_{0j}$  at the second level. The ICC can be important in multilevel analyses because it can allow determining the extent to which principals' responses vary among states and to which teachers' responses vary among schools (Raudenbush & Bryk, 1992).

Next, to check the first and second hypotheses, I set research models, the intercepts as outcome model in which level 1 intercept could be explained by the level 2 predictors (Hofmann, Griffin, & Gavin, 2000). From this research model, I can confirm influential factors on principal's influence on instruction and their facilitation of teachers' learning. To check whether there is relationship between states' accountability strength and principals' professional development support or their influence on instruction, I added four types of variables model: accountability strength, teachers' individual variables, school variables, and school climate variables.

The states' accountability strength variables were the proficiency performance standards and AMO strength, and high school graduation exit tests. The proficiency performance standards were the sum of the proficiency performance standards z-scores for reading and math and AMO strength were the sum of AMO strength z-scores for reading and math.

In the hierarchical linear modeling, there are three “centering” options to help interpret results (Hofmann & Gavin, 1998; Raudenbush & Bryk, 1992): “raw score (no centering), grand mean centering (in which individual scores are deviated from the grand mean), and group mean centering (in which individual scores are deviated from their respective group means)” (Gavin & Hofmann, 2002, p. 28). Although the appropriate selection of centering depends on the research model, grand-mean centering generally provides better estimates and interpretability (Whitener, 2001). Based on these findings, I used grand mean centering for variables except for dummy variables in my research model. The intercept as outcome model is represented below.

Principal level: Influence on instruction<sub>ij</sub>,

Support of professional work<sub>ij</sub>,

Provision of professional days before the school year<sub>ij</sub>, or

Provision of professional days during the school year<sub>ij</sub>

$$\begin{aligned}
 &= B0 + B1*(Gender) + B2*(Race) + B3*(Educational background) + \\
 &\quad B4*(Years as principals) + B5*(ASPIRING programs) + \\
 &\quad B6*(Professional development participation) + B7*(Suburban) + B8*(Rural) + \\
 &\quad B9*(Size) + B10*(SES) + B11*(Teachers' shared responsibility) + \\
 &\quad B12*(Student learning attitude) + B13*(Schools' resource adequacy) + R
 \end{aligned}$$



State level:  $B_0 = G_{00} + G_{01} * (\text{The proficiency performance standards}) +$   
 $G_{02} * (\text{AMO strength}) + G_{03} * (\text{High school graduation exit exams}) + U_0$

$$B_1 = G_{10}$$

$$B_2 = G_{20}$$

$$B_3 = G_{30}$$

$$B_4 = G_{40}$$

$$B_5 = G_{50}$$

$$B_6 = G_{60}$$

$$B_7 = G_{70}$$

$$B_8 = G_{80}$$

$$B_9 = G_{90}$$

$$B_{10} = G_{100}$$

$$B_{11} = G_{110}$$

$$B_{12} = G_{120}$$

$$B_{13} = G_{130}$$

## 6. Analysis of Teachers' Responses to Accountability Policies

In order to investigate the third and forth research hypotheses, I used three-level hierarchical linear modeling (HLM). When using HLM analysis, we have to consider sample size (Bell, Morgan, Kromrey, & Ferron, 2010) because small sample and cluster size can produce biased and inaccurate estimates (Bell et al., 2010). Especially, using a large-scale data set such as SASS, researchers have experienced the difficulties of data sparseness: few individuals are dispersed among a large number of level-2 units (Bell, Ferron, & Kromrey, 2008).

Adequate sample size at each level for analysis designs can be adjusted based on different interests in “parameter estimates, different expectation of statistical power, and different ranges of tolerable bias and accuracy” (Shih, 2008, p. 93). A 30/30 rule (30 groups with 30 individuals) for relatively unbiased and accurate random component estimates is normal in educational researches (Maas & Hox, 2004). Concretely, to produce more valid estimates of level 1 intercept variance ( $\sigma^2$ ), level 2 intercept variance ( $\tau_{00}$ ), and the level 2 slope variance ( $\tau_{11}$ ), at least a group size of 5 (at least 100 groups), 10 (at least 100 groups), and 20 (at least 200 groups) is needed (P. Clarke & Wheaton, 2007). When you examine interactions across levels, a minimum of 20 observations (level-1) for 50 groups (level-2) is recommended (Hox, 1998).

For unbiased and efficient estimates of the fixed-effects and variance components, we need “10 observations per group (even at low ICC values) as long as there are at least 200 groups” (P. Clarke & Wheaton, 2007, p. 345). “If one is willing to accept a standard error that is 5% higher than this minimum, then cluster number can be as low as 9” (Snijders & Bosker, 2012, p. 186). However, because the number of groups is more important than group size to produce unbiased estimates (P. Clarke & Wheaton, 2007), when there are many numbers of groups, fixed effects were affected by small group size (Theall et al., 2011; Maas & Hox, 2002).

Based on these literature reviews, I modified a sample size. There may not be any problems to examine the first and second questions, because each state has sufficient number of schools: in SASS dataset, Hawaii has 23 schools and California has 95 schools. However, insufficient teacher respondents in each school can make difficulties analyzing the third and fourth research questions: the effects of accountability systems on teachers via principals' different responses. For examples, no state had a school that includes seventeen teacher respondents. Sixteen Florida schools had only one teacher responded and ten California schools had two teacher respondents (see detail Appendix E).

After considering these conditions, I decided to use information from schools in which seven teachers responded for 3-level HLM analysis. When I set a cluster size as 10, 9 or 8, I can use principals and teachers from only twenty-six, forty, or forty-six states can be examined. The analysis using a small number of states may be not meaningful to examine the research questions: the relationship between strength of accountability systems and teachers' responses. Rhode Island was excluded from analysis because three states do not have seven schools that have seven teacher respondents. Also, California, Nebraska, and Utah did not have proficiency performance standards; so three states also excluded. Therefore, to respond the third and fourth research questions, I analyzed teachers who come from school with minimum seven teacher respondents in 46 states: 10,840 teachers come from 1,198 schools in 46 states. Descriptive statistics for the 3-level analysis variables appear in Table III-8.

**Table III-8 Descriptive Statistics for the 3-level Analysis Variables**

	Variable name	N	Mean	SD	Min	Max
Level-1	Professional development time for content	10,840	2.03	1.43	0.00	4.00
	Professional development time for instruction	10,840	0.77	1.05	0.00	4.00
	Professional development time for classroom management	10,840	0.61	0.87	0.00	4.00
	Teacher curriculum autonomy	10,840	2.99	0.89	1.00	4.00
	Teacher instructional autonomy	10,840	3.68	0.40	1.00	4.00
	Gender	10,840	0.59	0.49	0.00	1.00
	Race	10,840	0.92	0.27	0.00	1.00
	Educational background	10,652	0.54	0.50	0.00	1.00
	Teaching years	10,840	14.40	11.55	-1.00	54.00
	High qualified teachers	10,840	0.87	0.33	0.00	1.00
	Eighth grade & test subject	10,840	0.03	0.17	0.00	1.00
Level-2	Influence on instruction	1,198	3.65	0.41	1.33	4.00
	Support of professional work	1,198	2.36	0.97	0.00	4.00
	Provision of professional days before school year	1,198	0.96	0.19	0.00	1.00
	Provision of professional days during school year	1,198	0.94	0.25	0.00	1.00
	Suburban	1,198	0.52	0.50	0.00	1.00
	Rural	1,198	0.25	0.43	0.00	1.00
	School size	1,198	4.39	0.98	1.00	5.00
	School SES	1,198	2.18	1.35	0.00	5.00
	Teachers' shared responsibility	1,198	2.93	0.30	1.94	3.71
	Student learning attitude	1,198	2.73	0.46	1.21	3.93
	Schools' resource adequacy	1,198	2.65	0.26	1.86	3.56
Level-3	The proficiency performance standards	46	-0.02	0.94	-2.46	2.66
	AMO strength	46	0.01	1.00	-1.89	2.45
	High school graduation exit exams	46	0.48	0.51	0.00	1.00

First, I set a fully unconditional model. This model allowed me to determine the extent to which teachers' responses varied among states. The fully unconditional model is represented below.

Teachers level:

$$\text{Teacher curriculum autonomy}_{ijk} = P0 + E$$

$$\text{Teacher instructional autonomy}_{ijk} = P0 + E$$

$$\text{Teachers' professional development time for content}_{ijk} = P0 + E$$

$$\text{Teachers' professional development time for instruction}_{ijk} = P0 + E$$

$$\text{Teachers' professional development time for classroom management}_{ijk} = P0 + E$$

$$\text{Principals level: } P0 = B00 + R0$$

$$\text{State level: } B00 = G000 + U00$$

Teacher curriculum autonomy<sub>ijk</sub> = The level of teacher curriculum autonomy of teacher i in school j in state k

Teacher instructional autonomy<sub>ijk</sub> = The level of teacher instructional autonomy of teacher i in school j in state k

Teachers' professional development time for content<sub>ijk</sub> = The level of teacher's professional development time related to content of teacher i in school j in state k

Teachers' professional development time for instruction<sub>ijk</sub> = The level of teacher's professional development time related to instruction of teacher i in school j in state k

Teachers' professional development time for classroom management<sub>ijk</sub> = The level of teacher's professional development time related to classroom management of teacher i in school j in state k

P0 = Each principals' mean of teacher autonomy for curriculum and instruction and teachers' professional development time for content, instruction, and classroom management

B00 = Each state's mean of teacher autonomy for curriculum and instruction and teachers' professional development time for content, instruction, and classroom management

G000 = Grand mean of teacher autonomy for curriculum and instruction and teachers' professional development time for content, instruction, and classroom management

E = The teacher level variance

R0 = The principal level variance

U00 = The state level variance

To study whether or not there are relations among states' accountability strength, changed principal' behaviors, and teachers' autonomy, I implemented three level HLM analyses. Research model for the third research question was as follows:

Teacher level: Teacher curriculum autonomy<sub>ijk</sub> or Teacher instructional autonomy<sub>ijk</sub>

$$= P0 + P1*(Gender) + P2*(Race) + P3*(Educational background) + P4*(Years as teachers) + P5*(High qualified teachers) + P6*(Eighth grade & Test subjects) + E$$

Principal level:  $P0 = B00 + B01*(Influence\ on\ instruction) + B02*(Suburban) +$

$$\begin{aligned}
& B03*(Rural) + B04*(Size) + B05*(SES) + \\
& B06*(Teachers' shared responsibility) + \\
& B07*(Student learning attitude) + \\
& B08*(Schools' resource adequacy) + R0
\end{aligned}$$

$$P1 = B10$$

$$P2 = B20$$

$$P3 = B30$$

$$P4 = B40$$

$$P5 = B50$$

$$P6 = B60$$

$$\begin{aligned}
\text{State level: } B00 = & G000 + G001* (\text{The proficiency performance standards}) + \\
& G002* (\text{AMO strength}) + \\
& G003* (\text{High school graduation exit exams}) + U00
\end{aligned}$$

$$B01 = G010$$

$$B02 = G020$$

$$B03 = G030$$

$$B04 = G040$$

$$B05 = G050$$

$$B06 = G060$$

$$B07 = G070$$

$$B08 = G080$$

$$B10 = G100$$

$$B20 = G200$$

$$B30 = G300$$

$$B40 = G400$$

$$B50 = G500$$

$$B60 = G600$$

To examine the relationship among states' accountability strength, changed principal' behaviors, and teachers' professional development participation time, I implemented three level HLM analyses. Research model for the fourth research question is as follows:

Teacher level:

Teachers' professional development time for content<sub>ijk</sub> or

Teachers' professional development time for instruction<sub>ijk</sub> or

Teachers' professional development time for classroom management<sub>ijk</sub>

$$= P0 + P1*(Gender) + P2*(Race) + P3*(Educational background) + \\ P4*(Years as teachers) + P5*(High qualified teachers) + \\ P6*(Eighth grade \& Test subjects) + E$$

Principal level:  $P0 = B00 + B01*(Support\ of\ professional\ work) +$

$B02*(Provision\ of\ professional\ days\ before\ the\ school\ year) +$

$B03*(Provision\ of\ professional\ days\ during\ the\ school\ year) +$

$B04*(Suburban) + B05*(Rural) + B06*(Size) + B07*(SES) +$

$B08*(Teachers' shared responsibility) +$



$$B09*(\text{Student learning attitude}) +$$

$$B010*(\text{Schools' resource adequacy}) + R0$$

$$P1 = B10$$

$$P2 = B20$$

$$P3 = B30$$

$$P4 = B40$$

$$P5 = B50$$

$$P6 = B60$$

$$\text{State level: } B00 = G000 + G001* (\text{The proficiency performance standards}) +$$

$$G002* (\text{AMO strength}) +$$

$$G003* (\text{High school graduation exit exams}) + U00$$

$$B01 = G010$$

$$B02 = G020$$

$$B03 = G030$$

$$B04 = G040$$

$$B05 = G050$$

$$B06 = G060$$

$$B07 = G070$$

$$B08 = G080$$

$$B09 = G090$$

$$B010 = G0100$$

$$B10 = G100$$

$$B20 = G200$$

$$B30 = G300$$

$$B40 = G400$$

$$B50 = G500$$

$$B60 = G600$$

## **7. Limitations**

Although this paper had several limitations, the biggest limitation was disregarding the effects of districts. Districts tend to have a power for allocating financial and human resources to schools and educational activities (Gamoran & Dreeben, 1986). Moreover, in the age of accountability, districts set up a coherent vision, increasing students' achievement, implement district-wide curriculum, and provide district-wide professional development programs for teachers to develop their teaching quality (Bae, 2008; Luschei & Christensen, 2008; Hamilton et al., 2007; Togneri & Anderson, 2003). Although the districts' own standards and their accountability forces can influence principals and teachers' responses to state accountability policies (Louis et al., 2010), I should exclude district questionnaire because there were few districts for HLM analysis.

Next limitation is the effects of assistant principals. In many schools, there are assistant principals and they may practically implement many school activities. Generally, assistant principals implement various tasks, such as executing external communication and connection, implementing school staffs' development, and managing curriculum, learning, and teaching. Especially, as accountability demands increase, the instructional leadership role can become a major task because the accountability systems emphasize the students' academic outcomes (Oleszewski, Shoho, & Barnett, 2012). Although these assistant principals' behaviors can affect principals and teachers' responses to state accountability policies, I did not use assistant principals variable because the SASS data does not include enough information about assistant principals.

## **CHAPTER FOUR**

### **IV. RESULTS**

This chapter will describe the project results, which come from two-level and three-level hierarchical linear modeling (HLM) analysis. The results will be illustrated sequentially: principals' responses to states' accountability systems and the teachers' responses to states' accountability systems.

#### **1. Principals' Responses to States' Accountability Systems**

Principals' influence on instruction, their facilitation of teacher learning, and their provision of professional days were considered as the principals' responses to states' accountability systems. This part will describe the level, characteristics, and influential factors of principals' responses.

##### **1) The level and characteristics of principals' responses**

Based on the HLM analysis, I obtained the level and characteristics of the principals' responses to the states' accountability strength. Principals' influence on instruction was 3.648, and their support of professional work was about 2.367. In consideration of the fact that the maximum points of two responses were the same as 4, principals perceived that they had more influence on instruction than in supporting teachers' professional work. Principals' provision of professional development before the school year was .964 and during the school year was .926. This means that principals provide more professional days before the school years than during the school years.

The variance of principals' influence on instruction was 0.173, and the states' variance was 0.002. There were few differences in how principals' perceive their influence on instruction among the states. Each principal and state differently supports professional work, so principals' variance on support of professional work was 0.947, and the states' variance was 0.041. Principals' provision of professional days before the school year also differed by principals: the principals' variance was 0.032. However, there were little differences among states in the professional days before the school year: the states' variance was 0.002. Provision of professional days during the school year also had similar patterns. Although principals' variance was 0.068, states' variance was 0.002. To sum up, each principal differently responded to states' accountability policies, and especially principals' support of professional work had more variations in principals' levels. However, there were few differences in principals' responses to accountability policies among the states, except for principals' support of professional work. Principals in any state had a similar influence on instruction and provided professional days before or during the school year, although principals' support professional work may be different from states' accountability policies.

The ICC showed similar results. The ICC of principals' influence on instruction was approximately 1.137%, which means that the states' power over principals' influence on instruction is about 1.137 %. Principals' influence on instruction can be affected by principals' individual characteristics rather than by states' educational conditions. However, because of the states' different characteristics, principals who have the same individual characteristics may have different levels of influence on instruction.

The ICC of principals' facilitating teacher learning (support of professional work and provision of professional days) was larger than the ICC of principals' influence on instruction.

Three values of ICC of principals' support of professional work, provision of professional days before the school year, and provision of professional days during the school year were 4.1%, 5.2%, and 3.2%. The influences of principals' individual factors and schools' factors were 95.9%, 94.8%, and 96.8%, and states' accountability policies' influence were 4.1%, 5.2%, and 3.2%. Although principals' support of professional work and provision of professional days were affected by principals and school factors more than by states' accountability policies, principals who have the same individual and school characteristics can implement facilitating teacher learning according to their state's dissimilar educational policies. These results appear in Table IV-1.

**Table IV-1 The Level of Characteristics of Principals' Responses**

Principals' responses	Coefficient	Standard Error	Variance		ICC
			Level-1	Level-2	
Influence on instruction	3.648	0.010	0.173	0.002	0.011
Support of professional work	2.367	0.035	0.947	0.041	0.041
Provision of professional days before the school year	0.964	0.007	0.032	0.002	0.052
Provision of professional days during the school year	0.926	0.009	0.068	0.002	0.032

After finding the level and characteristics of principals' responses to states' accountability policies, I examined the principals' influence on instruction and their facilitating teacher learning (support of professional development and provision of professional days) in 51 states. Principals in Illinois, Massachusetts, South Dakota, and New York tended to have higher influence on instruction, while Alaska, Maryland, and Michigan principals had low influence on instruction.

Principals in California, Maine, Illinois, and Texas were likely to implement supportive behaviors for teachers' professional learning, while Arkansas, Kentucky, and Michigan provided less support for teachers' professional learning. In fourteen states, including Pennsylvania and Washington, almost all principals provided professional days before the school year; however Indiana, New Jersey, Rhode Island, and Ohio provided fewer professional days before the school year. Five states, such as Pennsylvania, Iowa, and Delaware principals were likely to provide professional days during the school year, while Arizona, California, and Rhode Island principals may not. These results appear in Appendix F.

## **2) The relationship between the strength of states' accountability systems and principals' responses**

The two level HLM analysis is used to answer the first research question, about what the relationship between principals' responses and the states' accountability systems: the proficiency performance standards, the strength of annual measurable objectives (AMO), and high school graduation exit exams. This analysis can lead to check four hypotheses.

### ***Principals' influence on instruction***

Principals' influence on instruction was not related to states' accountability systems. AMO strength was related to principals' perception of their influence on instruction. Principals in the states with large differences between starting points and intermediate goals were likely to have lower influence on instruction than principals in the states with low AMO strength. However, the proficiency performance standards and high school graduation exit exams requirement did not affect principals' influence on instruction. There may be little significant

difference in principals' influence on instruction between the states with high proficiency performance standards and difficult high school graduation exit exams and the states with low standard and no high school graduation exit exams requirement.

Therefore, the first hypothesis (states' high proficient performance standards, AMO strength, and high school graduation exit exams will be negatively and significantly correlated with principals' influence on instruction) was partially supported. The results are shown Table IV-2.

**Table IV-2 The Influential Factors for Principals' Influence on Instruction**

Fixed Effect	Coefficient	S. E.
Principals' influence on instruction	3.508	0.087
<b>State level</b>		
The proficiency performance standards	0.014	0.012
AMO strength	-0.017*	0.009
High School graduation exit exams	-0.004	0.009
<b>Principal (school) level</b>		
Gender	-0.027	0.019
White	-0.035	0.027
Educational background	-0.010	0.012
The years as principal	-0.001	0.001
ASPIRING program	0.038*	0.016
Professional development participation	0.147*	0.067
Suburban	0.057*	0.026
Rural	0.034	0.026
School Size	-0.007	0.009
School SES	0.008	0.007
Teachers' shared responsibility	-0.006	0.026
Student learning attitude	0.004	0.017
Schools' resource adequacy	0.105**	0.025

\*\*\*  $P < 0.000$ , \*\*  $P < 0.010$ , \*  $P < 0.050$



This model also found principal influential and school factors related to principals' reports on the extent to which they influence instruction. Principals' influence on instruction is related to principals' participation in development programs for ASPIRING school principals, which are formal programs implemented by many school districts to increase principals' abilities and to have a pool of capable principals. Through ASPIRING programs, principals can improve their capacities that can then be effective to teachers' professional works (Corcoran, Schwartz, & Weinstein, 2012). Professional development programs encourage principals to acquire better understanding of students' academic outcomes and to establish a school climate that may be directly related to students' development (O'Donnell & White, 2005). Knowledge and information for instruction acquired by these formal professional developments can lead principals to have more influence on instruction.

Principals in suburban schools are more likely to have influence on instruction than principals in urban schools. Because suburban school students have been considered as having high academic achievements, school districts may have less concern about the principals' capacities to establish curriculum, to set performance standards, and to evaluate teachers (Bloom & Owens, 2013), and principals in these suburban schools may feel less pressure from accountability policies. Urban schools may also have trouble with shortages of capable principals (Owings, Kaplan, & Chappell, 2011). Competent principals may choose suburban schools; so suburban school principals can have more influence on instruction than urban school principals. However, there was no significant difference in principals' influence on instruction in rural schools relative to urban schools.

Schools' resource adequacy was related to principals' influence on instruction. Principals

in schools with enough salary, sufficient educational materials, and low paperwork may have more influence on instruction than other principals. In schools with ample educational resources that can provide effective instruction, principals consider themselves as valuable instructional leaders (Spiri, 2001), and thus they can more have influence on instruction.

However, other factors such as principals' gender, their race, educational background, years as principals, school size, school SES, teachers' shared responsibility, and students learning attitude, did not affect principals' influence on instruction.

### ***Support of professional work***

The two level HLM examined the second hypothesis. Principals' support of professional work was not related to three states' accountability systems; the proficiency performance standard, AMO strength, and high school graduation exit exams. There may be little differences in principals' support of professional work in states based on the varying elements of accountability systems. Therefore, the second hypothesis (states' high proficient performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with principals' facilitation of teachers learning) was not supported.

This two level HLM analysis can identify the influential factors for principals' support of professional work. Principals' educational background, years as principals, ASPIRING programs, teachers' shared responsibilities, and schools' resource adequacy were significant factors rather than the strength of states' accountability systems. Principals' educational background and teaching years can increase principals' support for professional work. Principals with a high educational degree, such as specialist or doctoral degree, may provide more support for teachers' professional work than other principals. The years as principals appears to increase principals'

support of professional work. Novice principals may have insufficient knowledge about the technical aspects of school leadership and limited understanding of human relationships (Nelson, De la Colina, & Boone, 2008). Lack of knowledge and experience can lead to less support of professional work.

In addition, ASPIRING programs can enhance principals' support for professional work. Principals who participate in development programs for ASPIRING school principals can support teachers' professional work better than principals who did not participate in these programs. The ASPIRING programs can develop personal and professional qualities and behaviors that are related to teachers' professional work and school effectiveness (Corcoran et al., 2012). The differences of knowledge can substantially shape how principals led the work and responded to accountability policies (Louis & Robinson, 2012).

School climate can affect principals' support for professional work. Teachers' shared responsibilities in each school can be positively related to principals' support for professional work. Principals in schools where teachers own high responsibility for students' academic outcomes may provide support for teachers' professional work, including reducing teacher work loads and offering substitute teachers. Because principals may know about their teachers' work and what is required for high performance, they extend more effort to support their staffs' professional work.

However, principals' gender, race, and professional development participation, and schools' region, size, and SES did not affect principals' support of professional work. Schools' resource adequacy was also not a significant factor for principals' support of professional work.

### *Provision of professional days before or during the school year*

From the two level HLM analyses, we can examine the third and fourth hypothesis. Among three states' accountability systems, high school graduation exit exams were important influential factors for provision of professional days before the school year and during the school year. The principals in states with high school graduation exit exams may provide fewer professional days before and during the school year. The literature indicates that high stakes tests tend to narrow the curriculum for disadvantaged students, to focus on test-taking skills, and to decrease instruction time for untested subjects (Gayler, 2005). Principals in states with high school graduation exit exams receive pressure encouraging higher student pass rates on the tests. This stress may make principals focus more on students' learning, such as by implementing mandatory test previews and reviews classes (Holme, 2008). Moreover, because high stakes tests may emphasize basic skills, principals may not feel the necessity to provide professional days for improving teachers' capabilities. Therefore, high school graduation exit exams may be negatively associated with principals' provision of professional days before and during the school year. However, the proficiency performance standard and AMO strength did not affect principals' provision of professional days before and during the school year. Therefore, the third and fourth hypotheses were not supported. The results are shown Table IV-3.

Principals' provision of professional days before the school year was influenced by no principal and school characteristics. Regardless of principals' and schools' factors, principals provided professional days before the school year. Only high school graduation exit exams can influence principals' provision of professional days before the school year.

**Table IV-3 The Influential Factors for Principals' Facilitating Teacher Learning**

Fixed Effect	Facilitating teacher learning					
	Support of professional work		Provision of professional day before the school year		Provision of professional day during the school year	
	Coeff.	S. E.	Coeff.	S. E.	Coeff.	S. E.
Principals responses	1.907	0.205	0.949	0.039	0.799	0.079
<b>State level</b>						
The proficiency performance standard	0.018	0.041	-0.012	0.007	0.006	0.009
AMO strength	0.038	0.031	0.013	0.007	-0.007	0.008
High School graduation exit exams	0.003	0.065	-0.033 *	0.015	-0.034 *	0.017
<b>Principal (school) level</b>						
Gender	-0.063	0.050	-0.005	0.009	0.017	0.012
White	-0.081	0.085	-0.009	0.010	-0.012	0.019
Educational background	0.097 **	0.034	-0.004	0.005	-0.003	0.008
The years as principal	0.007 *	0.003	0.000	0.001	0.001	0.001
ASPIRING program	0.145 ***	0.035	0.009	0.008	-0.011	0.010
Professional development participation	0.260	0.153	0.022	0.028	0.094	0.068
Suburban	-0.090	0.066	-0.013	0.010	0.016	0.016
Rural	-0.113	0.070	-0.004	0.010	0.012	0.018
School Size	0.011	0.016	0.004	0.003	0.005 **	0.004
School SES	0.017	0.014	0.003	0.003	0.010 **	0.004
Teachers' shared responsibility	0.104 *	0.051	-0.006	0.013	0.049 *	0.019
Student learning attitude	-0.102 *	0.044	-0.009	0.010	-0.024 *	0.011
Schools' resource adequacy	0.036	0.079	0.011	0.015	0.000	0.021

\*\*\*  $P < 0.000$ , \*\*  $P < 0.010$ , \*  $P < 0.050$

Principals' provision of professional days during the school year was influenced by few principal and school characteristics. In schools with large size and with high SES students, principals can provide more professional days during the school. Teachers' shared

responsibilities encouraged principals to provide professional days during the school year. Principals who recognize their teachers' high shared responsibilities may believe that additional professional programs can be useful to increase students' academic achievement, and that their teachers may actively participate in these professional days. Positive student learning attitude may decrease principals' support for professional work. Principals in the schools with positive students' learning attitude may not feel the necessity for professional days during the school year. However, other factors affect principals' provision of professional days during the school year. The results are shown Table IV-3.

### *Synthesis of principals' responses*

Among states' accountability systems, AMO strength and high school graduation exit exams were negatively related to principals' responses: AMO strength may decrease principals' supporting for teacher learning and high school graduation exit exams can reduce principals' provision of professional days before and during the school year. However, proficiency performance standards did not affect four types of principals' responses. Based on these results, the first hypothesis (states' high proficient performance standards, AMO strength, and high school graduation exit exams will be negatively and significantly correlated with principals' influence on instruction) was partially supported. However, the second hypothesis (states' high proficient performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with principals' support of professional work), the third hypothesis (states' high proficient performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with principals' provision of professional days before the school year), and the fourth hypothesis (states' high proficient

performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with principals' provision of professional days during the school year) were not supported.

Principals' influence on instruction and their support for teacher learning were affected by principals' individual factors although principals' provision of professional days before and during the school year were not. School climate had an effect on principals' support for teacher learning and their provision of professional days during the school year.

## **2. Results of Teachers' Response to States' Accountability Systems**

Teachers' responses to states' accountability systems include teacher autonomy for curriculum and instruction and principals' participation time in programs related to content, instruction, and classroom management. This part will describe the level, characteristics, and influential factors of teachers' responses.

### **1) The level and characteristics of teachers' responses**

The 3-level HLM analyses enabled me to obtain the level and characteristics of the responses of principals and teachers. Teachers had two types of teacher autonomy: curriculum autonomy and instructional autonomy. The value of teacher curriculum autonomy was 2.989, and the value of teacher instructional autonomy was 3.672. These results show that teachers had more autonomy for evaluating and grading students, disciplining students, and determining the amount of homework to be assigned, than autonomy for selecting textbooks and other instructional materials, selecting content, topics, and skills to be taught, and selecting teaching techniques.

In teacher curriculum autonomy, teacher variance, school variance, and state variance was 0.681, 0.058, and 0.052. Although there were huge variances of teacher curriculum autonomy among teachers, there were few differences in school and states variances. Teacher variance of teacher instructional autonomy was 0.151 and school variance and state variances were 0.008 and 0.002. This means that each school and state may have similar teacher instructional autonomy, while teachers require different perceptions.

These results show the value of the ICC. 3-level HLM analysis has two ICC: 2-level ICC, which the proportion of school-level variance of the total variance, and 3-level ICC, which



means the proportion of state-level variance of the total variance. The 2-level ICC of teacher curriculum autonomy was approximately 7.3%, and the 3-level ICC was 6.6%. This means that when teachers implement autonomy, the influence of principal and school characteristics was 7.3%, and the power of states' accountability policies was 6.6% although the effects of teachers' individual factors was 86.1%. Although teachers may be more influenced by features of their individual factors than by schools', principals', and states' features, they can have different levels of teacher autonomy based on their principals', schools', and states' characteristics.

The 2-level ICC of teacher instructional autonomy was approximately 5.0%, and the 3-level ICC was 1.2%. When teachers implemented instructional autonomy, the power of schools factors and states accountability policies were 5.0% and 1.2%. Although the influence of schools and states on teacher instructional autonomy may not be bigger than the influence of teachers' characteristics, principals', schools', and states' characteristics can make a difference in teacher instructional autonomy.

Based on these results, teachers have more instructional autonomy than curriculum autonomy. The variance of teacher autonomy is different based on field: the variance of teacher curriculum autonomy was bigger than the variance the teacher instructional autonomy. In teacher curriculum autonomy, school level variance and state level variance were almost the same, so there may be little difference in teacher instructional autonomy among schools and states.

Teachers' participation time in professional development varied in professional development programs. Teachers spent more time participating in content programs (2.046) than in instruction programs (0.775) or in classroom management programs (0.600). This means that teachers may have spent almost 9-16 hours in the past 3 years participating in content programs, and they may have spent less 8 hours in the past 3 years on professional development related to

instruction and classroom management.

There is a lot of teacher variance among teachers' participation time in professional development programs. Teachers' variance in content programs was big, as 1.947. Some teachers may spend more time for content programs participation, but other teachers may not. However, the teacher variance in instruction program participation and classroom management program participation was 0.889 and 0.698.

There were also big differences in teachers' participation time in professional development programs among schools. Especially, principals in some schools may spend more time in instruction programs than principals in other schools: the school level variance of instruction programs was 0.146. When teachers spend time on participation in professional development, they may be influenced by school features. However, there may be a few differences of content and classroom management program participation time among schools (0.053 and 0.038).

In the state level variance, three types of teacher professional development participation time had similar value: variances of content, instruction, and classroom management were 0.047, 0.059, and 0.014. This means that teachers in 50 states may spend similar participation time in professional development programs.

The 2-level ICC of teachers' participation in professional development programs related to content, instruction, and classroom management were 2.6%, 13.3%, and 5.1%. Although the power of schools and principals characteristics on content and classroom management programs was low, teachers were influenced by the schools and principals features for their participation in instruction programs.

The 3-level ICC of professional development time for content, instruction, and classroom

management was 2.3%, 5.4%, and 1.9%. Teachers' spending time in professional development may be more influenced by features of their states' accountability policies, although the influences were smaller than the influence of teachers' individual characteristics. Although teachers have the same characteristics, states' accountability policies can have different time for professional development. These results appear in Table IV-4.

**Table IV-4 The Level of the Characteristics of Teachers' Responses**

Teachers' reponses	Coeff.	Standard Error	Variance			ICC 2-level	ICC 3-level
			Level-1	Level-2	Level-3		
Teacher autonomy							
- Curriculum	2.989	0.036	0.681	0.058	0.052	0.073	0.066
- Instruction	3.672	0.008	0.151	0.008	0.002	0.050	0.012
Participation time in PD							
- Content	2.046	0.036	1.947	0.053	0.047	0.026	0.023
- Instruction	0.775	0.039	0.889	0.146	0.059	0.133	0.054
- Classroom management	0.600	0.021	0.698	0.038	0.014	0.051	0.019

After understanding the level and characteristics of teacher autonomy and their participation time in professional development programs, I analyzed two types of teachers' behaviors by in states. North Dakota, Iowa, and Minnesota teachers may have high teacher curriculum autonomy and teacher instructional autonomy. Texas, Maryland, and Virginia teachers may have lower teacher curriculum autonomy and instructional autonomy than other states' teachers. However, Each state may have similar levels of teacher autonomy based on the types.

Teachers in Arkansas, Utah, Texas, and Vermont may spend more time on professional development programs related to the content than teachers in Indiana, New Jersey, and Mississippi do. Florida, Oregon, and Iowa teachers can spend more time on professional development programs related to instruction than New Jersey, Georgia, and Oklahoma teachers.

Arkansas, Texas, and Tennessee teachers were likely to join classroom management professional development programs but Maine, Connecticut, and New Mexico teachers may participate little in professional development about classroom management. Based on these results, Texas and Arkansas teachers may spend more time on professional development programs, while Connecticut and New Mexico teachers are less likely to participate in professional development related to content and classroom management. Teachers in 47 states may have different participation time by the three types of professional development programs. These results appear in Appendix G.

## **2) The relationship between states' accountability strength and teacher autonomy**

The three-level HLM analysis can answer the second research question, about what the relationship between the strength of states' accountability systems and teachers' responses is, and can check the fifth hypothesis (states' high proficient performance standards, AMO strength, and high school graduation exit exams will be negatively and significantly correlated with teacher curriculum autonomy) and the sixth hypothesis (states' high proficient performance standards, AMO strength, and high school graduation exit exams will be negatively and significantly correlated with teacher instructional autonomy)

### ***Teacher curriculum autonomy***

The fifth hypothesis can be checked by the three-level HLM analysis. The results showed that states' high proficiency performance standards significantly and positively influenced teacher autonomy related to curriculum. The proficiency performance standards can be relatively long term goals that schools should acquire by 2012. In the 2007-2008 school years when the

survey was implemented, teachers might have considered these standards as clear targets for making curricular choices and motivation, not as pressure. Therefore teachers in states with high proficiency performance standards enhanced teachers' sense of autonomy in the curriculum.

However, AMO strength were negatively related to teacher autonomy in the curriculum at the .100 significant level and high school graduation exit exams may be negatively related to teacher autonomy in the curriculum at the .050 significant level. AMO strength and high school graduation exit exams, as relatively short-term goals, perhaps produce more pressure than the proficiency performance standards, and thus two factors of states' accountability systems can decrease teachers' autonomy for selecting content and instructional materials. The results are shown Table IV-5.

Contrary to expectations that states' high proficient performance standards, AMO strength, and high school graduation exit exams will be negatively and significantly correlated with teacher autonomy for curriculum, the results of this study show that the accountability systems in the America send mixed signals to teachers to guide their work: the proficiency performance standards were positively related to teacher curriculum autonomy but AMO strength and high school graduation exit exams reduced teacher curriculum autonomy.

Principals' perceived influence on instruction is related to teacher autonomy for curriculum. Teachers reported more autonomous decisions about contents, textbooks, topics, and skills when their principals reported more power over instruction. Principals can be considered as protectors from the states' accountability systems, so principals' large influence on instruction can enhance teacher curriculum autonomy (Byrne, 2009; Crocco & Costigan, 2007).

**Table IV-5 Influential Factors for Teacher Curriculum Autonomy**

Fixed Effect	Coefficient	Standard Error
Teacher curriculum autonomy	2.189	0.150
<b>State level</b>		
The performance standard	0.074 *	0.029
AMO strength	-0.057 +	0.031
High School graduation exit exams	-0.113 *	0.056
<b>Principal (school) level</b>		
Principal's influence	0.080 ***	0.019
Suburban	0.096 **	0.034
Rural	0.191 ***	0.037
School Size	-0.093 ***	0.014
School SES	0.023 *	0.009
Teachers' shared responsibility	-0.021	0.040
Student learning attitude	0.027	0.026
Schools' resource adequacy	0.240 ***	0.056
<b>Teacher level</b>		
Gender	-0.018	0.015
Race	-0.046	0.044
Educational background	0.055 ***	0.014
Teaching years	0.014 ***	0.001
HQT	0.026	0.025
Eighth grade & test subjects	-0.390 ***	0.044

\*\*\*  $P < 0.000$ , \*\*  $P < 0.010$ , \*  $P < 0.050$ , +  $P < 0.100$

For teacher autonomy about curriculum, the school factors, such as their regions, size, and SES were significant factors. Teachers in suburban and rural schools have higher teacher autonomy for curriculum than teachers in urban schools. Perhaps because rural principals may perceive school staff as being involved in many decision making processes (Brown, Carr, Perry, & McIntire, 1996), rural school teachers report higher teacher autonomy. Moreover, because

urban schools are likely to have more low-performing students than suburban schools, teachers in urban schools might feel more pressure to meet their state's AYP standards (Sunderman, Orfield, & Kim, 2006). This pressure can lead urban school teachers to have lower curriculum autonomy than rural and suburban school teachers.

Schools' size was also a negative influential factor for teacher curriculum autonomy. Teachers in large schools tend to have lower teacher autonomy for curriculum than teachers in small schools. In small schools, teachers have more intimate and personal interactions with students (V. E. Lee & Loeb, 2000), and thus they can teach based on their students' needs, not by following the federal curriculum and its standards.

Teachers in school with high SES tend to have more teacher curriculum autonomy than teacher in school with low SES. School with low SES can have low performing students, who are not able to acquire states' proficiency performance standards. With this reason, teachers in low SES schools can try to follow states' standards and curriculum and can diminish their curriculum autonomy.

Among school climate factors, schools' resource adequacy influenced teacher autonomy for curriculum. Teachers in schools with resource adequacy were likely to have more teacher autonomy related to curriculum. High satisfaction and low paperwork, factors of resource adequacy, were considered as a significant factor of teacher autonomy. Teachers who manage their tasks and have lighter paper work may recognize themselves as being more autonomous (Pearson, 1995). Therefore, in schools with resource adequacy, teachers can have more control of curriculum. However, teachers' shared responsibilities and students' learning attitude were not significantly related to teacher curriculum autonomy.

Individual teacher characteristics, such as teachers' educational background, teaching

years, and their teaching grade and subjects, also influenced teacher curriculum autonomy. Because educational programs and teaching experiences can provide more knowledge related to curriculum, teachers who have more education background and who have long teaching years may have more teacher autonomy in curriculum decisions. These effects are significantly lower for eighth grade test-subject teachers. They may have little impact on decisions about the selection of textbooks and content because they have to teach a narrowed curriculum in order to produce high student test scores. The results are shown Table IV-5.

### ***Teacher instructional autonomy***

The three-level HLM analysis enabled to check the sixth hypothesis: whether states' high proficiency performance standards, AMO strength, and difficult high school graduation exit exams will be negatively and significantly correlated with teacher instructional autonomy. No states' accountability policies were a significant factor for teacher instructional autonomy. States' proficiency performance, AMO strength, and high school graduation exit exams did not influence teacher instructional autonomy. Because states' accountability policies focus on standards and curriculum rather than on instruction, in order to increase students' academic achievement (Diamond, 2012; Spillane et al., 2011), teachers may maintain their autonomy in instructional fields, including selecting teaching techniques, evaluating students, making decisions about homework, and disciplining students. Therefore, the sixth hypothesis was not supported. The results are shown Table IV-6.



**Table IV-6 Influential Factors for Teacher Instructional Autonomy**

Fixed Effect	Coefficient	Standard Error
Teacher instructional autonomy	3.557	0.075
<b>State level</b>		
The performance standard	-0.003	0.008
AMO strength	0.001	0.009
High School graduation exit exams	-0.005	0.009
<b>Principal (school) level</b>		
Principal's influence	0.021 *	0.012
Suburban	0.009	0.013
Rural	0.022	0.015
School Size	-0.004	0.006
School SES	0.006 <sup>+</sup>	0.005
Teachers' shared responsibility	0.079 **	0.022
Student learning attitude	0.046 ***	0.013
Schools' resource adequacy	0.090 ***	0.022
<b>Teacher level</b>		
Gender	0.048 **	0.008
Race	0.016	0.017
Educational background	-0.007	0.009
Teaching years	0.001 **	0.000
HQT	0.011 <sup>+</sup>	0.014
Eighth grade & test subjects	-0.053	0.010

\*\*\*  $P < 0.000$ , \*\*  $P < 0.010$ , \*  $P < 0.050$ , <sup>+</sup>  $P < 0.100$

Principals' influence on instruction can be positively related to teacher instructional autonomy. When their principals report that they hold more power over instruction, teachers perceive their principals as a protector from the states' accountability systems. Therefore, principals' influence on instruction can increase teacher instructional autonomy (Byrne, 2009; Crocco & Costigan, 2007).

Among school characteristics, only school SES influenced teacher instructional autonomy.

Teachers in schools with high SES students reported that they have more autonomous decisions about teaching techniques, disciplining students, and determining homework than teacher in schools with low SES students. However, other variables such as region and size did not affect teacher instructional autonomy.

Schools' resource adequacy, teachers' shared responsibility, and positive student learning attitude were crucial factors for teacher instructional autonomy. When there is a healthy school climate, which promotes teachers' collaboration, communication, and job satisfaction (Garvin, 2007; Pearson, 1995), teachers are likely to enhance teacher instructional autonomy (Sparks, 2012; Erpelding, 1999). Teachers in schools with high teachers' shared responsibility, positive students' learning attitudes, and sufficient school resources report higher teacher instructional autonomy than other teachers. However, no school physical factors were related to teacher instructional autonomy.

Among school individual characteristics, gender, teaching years, and highly qualified teachers were essential factors for teacher instructional autonomy. Female teachers reported more teacher instructional autonomy because female teachers prefer to enjoy school professional communities more than male teachers (Louis, Marks, & Sharon, 1996). Experienced teachers can implement autonomous decisions related to instruction, because novice teachers receive much more supervision than veteran teachers, and the supervision tends to be directive (Range, Scherz, Holt, & Young, 2011). Highly qualified teachers had more instructional autonomy than non-qualified teachers. The results are shown Table IV-6.

### ***Synthesis of teacher autonomy***

States' accountability systems significantly affected teacher curriculum autonomy but not

teacher instructional autonomy. For teacher curriculum autonomy, proficiency performance standards showed positive effects and AMO strength and high school graduation exit exams made negative effectives. The influence of states' accountability systems on teacher curriculum autonomy was mixed. However, states' accountability systems did not affect teacher instructional autonomy.

Based on these results, the fifth hypothesis (states' high proficient performance standards, AMO strength, and high school graduation exit exams will be negatively and significantly correlated with teacher curriculum autonomy) was partially supported, and the sixth hypothesis (States' high proficient performance standards, AMO strength, and high school graduation exit exams will be negatively and significantly correlated with teacher instructional autonomy) was not supported.

Principals' influence on instruction was positively related to teacher two types of autonomy. When principals reported more influence on instruction and provide sufficient resources, teachers had more power to make decisions about curriculum and instruction. School characteristics were significant for teacher curriculum autonomy, and school climate significantly affect teacher instructional autonomy. Experienced teachers also can more teacher autonomy in curriculum and instruction fields.

### **3) The relationship between states' accountability strength and teachers' participation time in professional development**

Through three-level HLM analysis, I can check the seventh hypothesis (states' high proficient performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with teachers' participation time in professional development programs related to content), the eighth hypothesis (states' high proficient performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with teachers' participation time in professional development programs related to instruction), and the ninth hypothesis (states' high proficient performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with teachers' participation time in professional development programs related to classroom management).

#### ***Content professional development participation time***

The analysis examined the seventh hypothesis: states' high proficiency performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with teachers' participation time in content professional development programs. States' high proficiency performance standards were associated to teachers' participation time in professional development programs about the content. The proficiency performance standards are goals that students should acquire by 2012. To attain these goals, teachers need to devote their time to develop their knowledge and capacities through content professional development programs in the 2007-2008 school year. Thus, teachers in states with high proficiency performance standards may encourage teachers to participate in professional

development programs for content.

However, AMO strength was negatively related to teachers' participation time in content professional development programs. Unlike the proficiency performance standards, the annual measurable objectives are short-term goals that students should achieve in the 2007-2008 school year. In order to avoid sanctions, teachers may focus on students' academic improvement, not on their knowledge development. Therefore, teachers in states with high AMO strength appear to spend less time in content professional development programs. However, high school graduation exit exams were not significantly associated to teachers' spending time in content programs. Based on these results, the seventh hypothesis can be partially supported. The results are shown Table IV-7.

Among principals' behaviors, professional days built in before the school year were an effective method for teachers to participate in content professional development programs at the .100 significant level. During the school year, teachers may not have sufficient time to prepare their curriculum and to improve their knowledge. Therefore, teachers preferred professional days before the school year for purposes of improving their readiness to implement the required curriculum.

School characteristics, such as school region and size, were important factors for teachers' participation time for content. Rural schools are recognized as having limited educational resources in order to meet states' standards (Arnold, Newman, Gaddy, & Dean, 2005). Insufficient resources can produce few professional development programs, and limited opportunities for teachers to participate in professional development programs may cause rural school teachers' low participation rates. Although suburban school teachers also spent less time in participation time in professional development content programs than urban teachers, there

might be a different reason. Because students in suburban schools are considered as having high academic achievement, suburban school teachers may not feel the necessity to participate in professional development programs compared to urban school teachers.

**Table IV-7 Influential Factors for Teacher's Participation Time in Content Professional Development Programs**

Fixed Effect	Coefficient	Standard Error
Teachers' participation time in content professional development	1.400	0.222
<b>State level</b>		
The proficiency performance standards	0.081 *	0.031
AMO strength	-0.074 *	0.032
High School graduation exit exams	-0.071	0.065
<b>Principal (school) level</b>		
Principal's support of professional work	0.020	0.016
Principals' provision of professional days before the school year	0.161 <sup>+</sup>	0.087
Principals' provision of professional days during the school year	0.058	0.061
Suburban	-0.094 *	0.038
Rural	-0.201 ***	0.050
School Size	0.048 *	0.022
School SES	0.009	0.013
Teachers' shared responsibility	0.124 <sup>+</sup>	0.064
Student learning attitude	-0.059 <sup>+</sup>	0.034
Schools' resource adequacy	-0.045	0.061
<b>Teacher level</b>		
Gender	0.117 ***	0.032
Race	-0.123 *	0.047
Educational background	0.077 **	0.028
Teaching years	0.006 ***	0.001
HQT	0.117 **	0.037
Eighth grade & test subjects	0.194 *	0.096

\*\*\*  $P < 0.000$ , \*\*  $P < 0.010$ , \*  $P < 0.050$ , <sup>+</sup>  $P < 0.100$

Teachers in large schools spent more time in content program participation than teachers in small schools. The number of students in large schools may encourage these schools to create various and comprehensive programs to address students' needs (K. R. Stevenson, 2006; V. E. Lee & Loeb, 2000). Many students with different needs might lead teachers to participate in professional development content programs.

School climate significantly influenced teachers' participation time in content programs. Teachers in school with high-shared responsibility among teachers can participate in content professional development programs. Shared responsibility may encourage teacher to spend more time on content programs. However, students' positive learning attitude reduced teachers' spending time for content professional development programs. Teachers in schools with positive learning attitude may not need to participate in content professional development programs.

Teachers reported differential benefits based on their individual attributes. Teachers' participation time in professional development programs about content may differ according to their gender. Female teachers are more likely to engage in interactive professional development about content than male teachers. Female teachers are likely to be involved in school professional community (Louis et al., 1996) based on their effective communication skills (Tannen, 1991). Teachers' race was also a significant aspect. White teachers spent less in content professional development programs than non-white teachers. White teachers can be assigned to high quality schools due to non-alternative teacher certification (Kee, 2012; Shen, 1997), and they may not feel the necessity to participate in professional development programs.

Teachers with high educational background, many teaching years, high qualifications, and grade and subject tend to be involved in content professional development programs. Because these types of teachers feel the necessity for improving their teaching quality in order to

support students' academic outcomes (Jackson, 2006; Steffy, 2000), they may spend more time on content professional development programs. Highly qualified teachers and eighth grade English or mathematics teachers were also likely to participate in professional development about content.

### ***Instruction professional development participation time***

The eighth hypothesis, states' high proficiency performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with teachers' participation time in instruction professional development programs, was also examined. All aspects of states' accountability systems, proficiency performance standards, AMO strength, and high school graduation exit exams were unrelated to teachers' participation time in instruction professional development programs. As I saw with the autonomy analyses, accountability pressures did not appear to penetrate into the classroom in the same way they influence curricular decisions. Based on these results, the eighth hypothesis cannot be supported. The results are shown Table IV-8.

Principals' support of professional work and their provision of professional days before and during the school year can be associated with teachers' spending time in instruction professional development although the association was not significant. When principals offer substitute teachers, common planning time, reduced teacher workloads, and professional days, teachers may more easily attend the type of instructional professional development programs. The supportive environment for teacher learning can encourage teachers to spend more time for instructional professional development programs. These results support the kind of embedded professional development and collaborative work required for instructional improvement.



**Table IV-8 Influential Factors for Teacher's Participation Time in Instruction Professional Development Programs**

Fixed Effect	Coefficient	Standard Error
Teachers' participation time in instructional professional development	0.723	0.304
<b>State level</b>		
The proficiency performance standards	0.055	0.035
AMO strength	0.051	0.038
High School graduation exit exams	-0.113	0.072
<b>Principal (school) level</b>		
Principal's support of professional work	0.020	0.014
Principals' provision of professional days before the school year	0.029	0.077
Principals' provision of professional days during the school year	0.052	0.058
Suburban	-0.078 <sup>+</sup>	0.040
Rural	-0.060	0.046
School Size	0.020	0.023
School SES	-0.013	0.013
Teachers' shared responsibility	0.104	0.068
Student learning attitude	-0.065	0.040
Schools' resource adequacy	-0.067	0.053
<b>Teacher level</b>		
Gender	0.149 ***	0.019
Race	-0.126 *	0.056
Educational background	0.035 <sup>+</sup>	0.020
Teaching years	0.000	0.001
HQT	0.030	0.038
Eighth grade & test subjects	0.184 ***	0.048

\*\*\*  $P < 0.000$ , \*\*  $P < 0.010$ , \*  $P < 0.050$ , <sup>+</sup>  $P < 0.100$

Schools' characteristics were related to teachers' participation time in instruction programs. Teachers in suburban locations were less likely to join in activities to improve

instruction. Because students in suburban schools may have high academic outcomes, teachers in these schools cannot need instruction professional development programs. However, other school characteristics and school climate did not offer significant effects on teachers' participation time in professional development programs related to instruction.

Teachers' individual factors, such as gender, race, teaching years, and teaching grade and subjects, were significant factors to increase teachers' participation time in instruction programs. Female teachers were much more inclined to pursue this type of professional learning than are male teachers. Teachers' race also affected teachers' participation time in programs focusing on instruction. Minority teachers may come from alternative teacher certification programs (Kee, 2012; Shen, 1997), and they may be assigned to low quality schools with non-excellent school climate and low SES. Therefore, non-white teachers may feel the necessity of this type of professional development programs compared to non-white teachers. Experienced teachers were more likely to spend more time on instructional programs.

The eighth grade English and mathematics teachers spent more time on content and instruction professional development programs. Eighth grade is a tested grade and English and mathematics are test sub-subjects. Eighth grade English and mathematics teachers may feel accountability pressures most strongly, so they may try to increase their teaching quality through professional development. The results are shown Table IV-8.

### ***Classroom management professional development participation time***

The three analysis of the ninth hypothesis (states' high proficiency performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with teachers' participation time in professional development programs related to

classroom management) was studied. States' accountability policies, proficiency performance standards, AMO strength, and high school graduation exit exams, did not affect teachers' participation time in professional development programs related to classroom management. Teachers' inclination to develop their management skills was unrelated to any dimension of accountability because these states' accountability systems may focus on standards, not classroom management. Based on these results, the ninth hypothesis cannot be supported. The results are shown Table IV-9.

Principals' facilitating teacher learning may not be an effective method for teachers to participate in classroom management professional development programs. Principals' support of professional work and their provision of professional days before and during the school year did not affect teachers' participation time in professional development programs related to classroom management.

School SES among school physical characteristics significantly affected teachers' participation in classroom management programs. When the schools have many students who qualify for the federal free or reduced-price lunch programs, the teachers in these schools may spend more time in classroom management programs. Schools with significant numbers of economically disadvantaged children may find it difficult to acquire AYP due to low academic achievement (Foy, 2008). To overcome the weakness, the teachers focus on classroom management professional development programs.

**Table IV-9 Influential Factors for Teacher's Participation Time in Professional Development Related to Classroom Management**

Fixed Effect	Coefficient	Standard Error
Teachers' participation time in classroom management professional development	0.848	0.155
<b>State level</b>		
The proficiency performance standards	-0.014	0.021
AMO strength	0.002	0.015
High School graduation exit exams	0.017	0.037
<b>Principal (school) level</b>		
Principal's support of professional work	0.010	0.009
Principals' provision of professional days before the school year	0.051	0.043
Principals' provision of professional days during the school year	-0.040	0.035
Suburban	0.030	0.027
Rural	0.015	0.038
School Size	-0.013	0.015
School SES	-0.020 *	0.009
Teachers' shared responsibility	0.040	0.039
Student learning attitude	-0.153 ***	0.028
Schools' resource adequacy	0.120 *	0.057
<b>Teacher level</b>		
Gender	-0.010	0.012
Race	-0.148 ***	0.037
Educational background	-0.038 *	0.019
Teaching years	-0.002 **	0.001
HQT	-0.043	0.027
Eighth grade & test subjects	-0.007	0.053

\*\*\*  $P < 0.000$ , \*\*  $P < 0.010$ , \*  $P < 0.050$ , +  $P < 0.100$

Schools' resource adequacy can increase teachers' classroom management professional development, although students' learning attitudes is negatively related to teachers' spending

time on classroom management professional development. When students have positive learning attitudes, teachers have less need for this focus so principals provide these programs; conversely, in schools where students' attitudes are negative, principals support teachers to invest more in professional development to improve their classroom management skills. Resources are necessary for these programs, thus schools with more resources are likely to have more of these types of programs available. However, teachers' shared responsibility did not affect teachers' participation in any professional development programs.

Classroom management training was not attractive to minority teachers or those with extensive experience or MA degrees. White teachers spend more time in classroom management programs than non-white teachers. Because the master's course can provide knowledge about classroom management, teachers with high educational background may not feel the necessity to participate in classroom management programs while teachers without a master's degree need more professional development programs related to classroom management. Experienced teachers participated less in classroom management programs because they can learn classroom management skills during their long teaching years.

### ***Synthesis of teachers' participation time in professional development***

States' accountability systems affected only teachers' participation time in content programs, not instruction and classroom management programs. The influence on three factors of accountability was mixed. The proficiency performance standards significantly increased teachers' participation time in content programs while AMO strength decreased the time. The high school graduation exit exams did not significantly influence teachers' spending time in classroom management programs.

Based on these results, the seventh hypothesis (states' high proficient performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with teachers' participation time in professional development programs related to content) was partially supported. However, the eighth hypothesis (states' high proficient performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with teachers' participation time in professional development programs related to instruction) and the ninth hypothesis (states' high proficient performance standards, AMO strength, and high school graduation exit exams will be positively and significantly correlated with teachers' participation time in professional development programs related to classroom management) were not supported.

Principals are essential factors for teachers' participation in professional development programs. Professional days before the school year that principals provide can promote teachers' spending time in professional development programs related to content. Principals' facilitating teachers learning can increase teachers' spending time in professional development programs.

School characteristics made effects on teachers' participation time in content professional development programs, and school climate affected teachers' spending time on classroom management programs. However, teachers' professional development time related to instruction was not affected by school characteristics and school climate.

Teachers' race was essential factors for teachers' participation time in three types of professional development programs. White teachers spent less time in three types of professional development programs. Eighth grade teachers who teach English and mathematics devoted more time for professional development programs related to content and instruction.

## **CHAPTER FIVE**

### **V. DISCUSSION, IMPLICATIONS, AND CONCLUSION**

School staffs, principals, and teachers differently respond to each state's accountability system. Their responses may be the fundamental key to successful school education and students' outcomes (Louis et al., 2010; DeMoss, 2002). This dissertation represents an empirical test of whether states' accountability policies are related to principals' and teachers' responses to them. The results of this study revealed the extent of principals' and teachers' responses to accountability, and showed the range of influential factors of states, principals, schools, and teachers. In this concluding chapter, I first discuss the major findings of the study as principals' and teachers' responses to accountability systems. At the end of the chapter, I suggest several implications of the study for teachers, school leaders, policymakers, and educational researchers.

#### **1. Discussion**

##### **1) The weak negative relationship between states' accountability policies and principals' responses**

Recognizing the differences in accountability policies among 50 states, I assumed that these differences could cause dissimilar responses from principals. This study about the relationship between states' accountability policies and principals' responses showed that there is a negative relationship between the strength of states' accountability systems and principals' response. Principals in state with large differences of starting points and intermediate goals had low influence on instruction and principals in states with high school graduation exit exams requirement especially provide fewer professional days before the school year. Principals in

states with strong accountability systems are likely to narrow the curriculum, to emphasize test-taking skills, and to decrease instruction time for untested subjects (Gayler, 2005). Moreover, they provide additional preview and review classes to help many students pass the tests (Holme, 2008). These principals' behaviors focus on students, not teachers. Therefore, the principals in states with strong accountability systems have low influence on instruction and provide less professional days to teachers before the school year.

Other studies also show similar results, in which states' accountability may produce negative effects on principals' perceptions and behaviors. Under the accountability contexts, principals feel personal and professional pressure from their central office, community, and themselves (Knobl, 2010; Priolo, 2010). This pressure leads principals to focus on test subjects. Principals offer more courses or extra-curricular programs only to test subjects (Priolo, 2010; Spillane et al., 2002), and they redirect funds to these subjects (Lewis, 2010; Ladd & Zelli, 2002). Principals also force teachers to narrow the curriculum and to spend more time on teaching test-taking skills (Hollingworth, Dude, & Shepherd, 2010; Jones & Egley, 2010; Gardiner, Canfield-Davis, & Anderson, 2009).

However, the relationship between states' accountability policies and principals' responses may not strong. The first assumption of the weak relationship between states' accountability policies and principals' influence on instruction and their facilitating teacher learning is that states' accountability policies are external mandates which are "complex arrangement[s] of policies, created by actors and interests outside of schools, who are in position to reward and punish schools, aimed at impacting practices inside schools, and requiring reporting to diverse external audience" (Knapp & Feldman, 2012, p. 667). This complicated combination may not be educationally coherent and can create conflicts with school staffs



(Firestone & Shipps, 2005; O'Day, 2002). Therefore, states' accountability policies, as external accountability systems, may have limitations to answer any problem related to teaching and learning (J. B. Smith, Smith, & Bryk, 1998).

Another assumption is the influence of the district. Within a state, each district may have different levels of accountability policies (Firestone et al., 1998), which makes a dissimilar relationship between principals. Because district practices can determine the principals' efficacy and behaviors (Leithwood, Louis, & Anderson, 2012; Louis et al., 2010), when districts have strong policies and a supportive relationship with their principals, principals may adapt the states' accountability policies or integrate the policies with their pre-existing educational missions (Louis & Robinson, 2012).

The effect of the media on all principals can be one reason why there is little relationship between states' accountability policies and principals' responses. Since the implementation of NCLB, principals have watched and listened to the horror of test scores by print and visual media (Foy, 2008). Through these media, even principals who belong to states with weak accountability systems can understand and feel strong accountability policies.

The last assumption is time. Initially, principals may have negative perceptions about accountability systems because their responsibilities shift from school management to the school effectiveness based on students' test scores (Foy, 2008). However, time can allow a principal to accept accountability policies (Louis & Robinson, 2012). Since the implementation of NCLB, principals gradually have made sense of the accountability systems and consider the systems as their policies (Louis et al., 2005). Therefore, in the 2007-2008 school year when after five years of NCLB implementation, principals did not differently respond to states' accountability systems based on the strength of states' accountability systems.

## **2) The directly opposed effects of states' accountability policies on teachers' responses**

Assuming the diverse level of states' accountability policies, I tried to answer the second research question: what is the relationship between the strength of states' accountability systems and teachers' responses, that are teacher autonomy and their participation in professional development programs. The analysis for the second research question found interesting results, which the factors of states' accountability policies produced the directly opposed effects on teachers' responses. The proficiency performance standards increased teacher curriculum autonomy and teachers' participation time in content-based professional development programs, although high school graduation exit exams decreased their curriculum autonomy and AMO strength diminished teachers' spending time in content focused professional development programs.

According to the results of these research models, AMO strength and high school graduation exit exams caused negative effects. Teachers in states with a big difference between starting points and annual measurable objectives, and in states with rigorous high school graduation exit exams may have lower teacher autonomy for curriculum and spend less time in content-focused professional development programs than teachers in states without these two state accountability policies. Achievement targets make a difference. Accomplishment of AYP goals is a relatively immediate matter for both teachers and students. To avoid sanctions, students should acquire AYP goals and pass the exams, and teachers should help student to obtain high test scores. However, longer term goals revolve around implementation of curricular standards. For students' successful outcomes, teachers may give up their autonomy and follow the state's

standards and curriculum, and, thus, they can focus on students' learning rather than developing their own capacities. Therefore, AMO strength and high school graduation exit exams can provide negative effects on teachers' responses.

However, the proficiency performance standards were positively associated with teacher curriculum autonomy and their participation time in professional development programs related to content. The proficiency performance standards can be relatively long term goals that teachers should acquire by 2012. In the 2007-2008 school year, when the survey was implemented, teachers might not have felt any pressure to acquire the proficiency performance standards, thus, they could maintain and develop their teacher autonomy. In addition, the proficiency performance standards provided direction for teachers to promote their capabilities and their instruction. The motivation perhaps led to teachers' participation in professional development programs, especially on content.

Based on these results, proficiency performance standards may be positively related to teacher curriculum autonomy and teachers' participation time in content professional development programs. These findings suggest that the recent waiver policy that federal government implemented over the past few years could produce positive effects. (Davidson, Reback, Rockoff, & Schwartz, 2013). Because it would be impossible for all schools to reach proficiency performance standards goals by 2014, the federal Department of Education started permitting states' flexibility requests to alleviate the impending 100% proficiency deadline in 2011. As of March 2013, all states but Nebraska and Montana had submitted flexibility requests, and thirty-five of these requests have been approved. With the flexibility policies, the principals and teachers may gain additional time to improve their students' academic accomplishment. Having time on teachers' side can be a motivation and a goal, not pressure, for teachers.

Therefore, through the flexibility policies of NCLB, teachers can enhance their autonomous decisions about curriculum and their participation time in professional development programs related to content fields.

### **3) The limited effects of states' accountability policies on specific schools**

The states' accountability systems can be significantly and negatively related to schools with specific features, including urban, large and poor schools. This study found that urban schools, large size schools, and schools with low SES students tended to have low teacher curriculum autonomy and to spend more time on professional development time related to content, which might be negatively related to states' accountability systems. Teachers in schools with limited educational resources also report low teacher curriculum autonomy.

Schools in urban areas and schools with low SES students and a large size are likely to have many low-performing students. Low students' academic achievement may make teachers feel pressure from the states' accountability systems because under the states' accountability systems teachers can receive some sanctions when students do not accomplish states' academic goals. With this reason, teachers in these urban, large, and poor schools really may follow states' standards and content for the tests, and, thus, they feel that they have no autonomy. Closely related, teachers in these types of schools report spending more time on professional development programs related to content perhaps to confirm and understand test contents and to increase their students' academic achievement. Then teachers appear to sense pressures of accountability perspective more so than teachers in suburban, small size, and affluent schools that have high-performing students.

The results suggest that low teacher curriculum autonomy might aggravate the

educational circumstance of urban, large, and poor schools even though it increases the teachers' participation in professional development that might be positive for high teacher quality. Teacher autonomy can be considered as essential source of teacher's intrinsic motivation, professionalism, and job satisfaction (Roth et al., 2007; Pearson & Moomaw, 2005). Teachers with sufficient autonomy can implement effective classroom instruction and have satisfaction, which can lead to retain in their teaching jobs. Under the accountability systems, teachers in urban, large, and poor schools appear to have low teacher curriculum autonomy, which can make teachers feel less impelled to participate in collaborative work, take a less professional perspective of their work, and be less willing to work on improving their teaching practice. Moreover, job dissatisfaction based on low teacher curriculum autonomy from states' accountability might lead to increased turnover of any capable teachers in schools with a poor educational environment. Although states' accountability systems intended to increase the academic achievement of low-income, low achieving, and minority students, these accountability systems might actually interrupt students' improvement in urban schools, in poor schools, and in large schools as a result of low teacher curriculum autonomy.

#### **4) The limited effects of states' accountability policies on specific domain of practice**

One more meaningful point is that the influence of states' accountability policies on teachers is limited to specific domains of practice. This study found that states' accountability policies did not affect teacher instructional autonomy and teachers' participation time in professional development programs related to instruction and classroom management.

Teachers' specific task domains of practice can be perceived by teachers in very different

ways within accountability contexts. Relatively, teacher instructional autonomy and teachers' participation in classroom management may be remote domains of practice for the states' accountability systems, because the goal of accountability policy may be to constrain the individual decisions teachers make in deciding what curriculum to follow in their practice. Under the accountability contexts, teachers have limited control about content and curriculum (Eden, 2001), and they devote their time to check and understand the content of tests. However, teachers appear to retain autonomy in how to teach (Desimone, 2013; Diamond, 2012; Spillane et al., 2011), and thus they may not feel the necessity for spending time on professional development programs related to instruction. Moreover, because teachers' classroom management may be more related to school contexts, such as school SES and student learning attitudes than states' accountability policies, teachers' participation time in professional development related to classroom management may not be affected by states' accountability policies.

The limited effects of states' accountability policies on teachers' specific fields can be also explained as recoupling, which "the process of creating tight couplings where loose couplings were once in place" (Hallett, 2010, p. 54). School organizations have been considered as loosely coupled systems, in which the external environment and policies may have rarely penetrated the instruction in classroom (Fullan, 2001). However, school organization may be comprised of two parts: the institutional sectors, in which loose coupling predominates and the technical sectors, where tighter coupling occurs (W. R. Scott & Meyer, 1983). Therefore, under the accountability systems, teachers may enjoy more autonomy for instructional decisions rather than curriculum decisions.

School climate results from the interaction of various people over time. How teachers work together and the extent to which they share responsibility for conditions outside the

classroom can influence school climate. Whether students come to school ready to learn or not, students contribute to general condition in the schools. Finally, the extent to which parents and communities support the school with adequate resources is related to the climate within the school.

With these reasons, school climate is an essential factor for teacher instructional autonomy and for teachers' participation in classroom management programs in this study. School climate can be effective to construct a collective sense-making process in schools (Louis et al., 2005), and so teachers in the schools with healthy school climate are more likely to collaborate and communicate each other (Garvin, 2007; Pearson, 1995). Through this interaction process, teachers can share not only school visions but also various educational knowledge and information. Therefore, positive school climate can encourage teachers to make autonomous decisions about instruction and classroom management (Finnigan, 2012; Sparks, 2012; Erpelding, 1999), which can be essential for school education improvement. It is also possible that teachers who work with other closely reach collective decisions for which they feel individual responsibility and control.

### **5) Principals' effects on teachers' responses**

Through this study, I found that principals' responses are essential factors for teachers' responses. Principals' facilitating teacher learning encourages teachers to spend more time in professional development programs, and the preferred modality and the timing of professional development varies based on the focus of the activities. Teachers' participation in professional development focusing on content is enhanced when principals offered the professional days before the school year. Content may require attention and planning before the school year begin.

In contrast, instruction is the center of teacher's daily work and thus teachers' learning for instruction can be enhanced by the principals' support of professional work during the school year. Principals' support, such as providing for substitutes, arranging for early dismissal, providing common planning time, and reducing teacher work loads, create a school environment which encourages teachers' professional growth development (Drago-Severson, 2012; Croft et al., 2010).

In addition, principals' reports of the extent to which they influence instruction show a positive association with teacher autonomy for curriculum and instruction. Based on the "win-win-game concept," principals' influence on instruction can have a positive relationship with teachers' power in decision-making (Shen & Xia, 2012). Under the accountability policies, school staffs, both principals and teachers, may be affected by pressure of states and districts. This pressure may produce a close identity between teachers and their principals. Teachers consider principals as protectors against the pressure of the state administration, and as producers of the school environment, which teachers need to implement their autonomy (Byrne, 2009; Crocco & Costigan, 2007). Therefore, principals' influence on instruction can encourage teacher autonomy.



## **2. Implications**

Based on the results and discussion of this study, it is clear that external accountability systems measured by this study do not enhance principals' instructional work or teachers' sense of control over their work on classroom condition generally. I would like to suggest several implications. First, recognizing the limitations of external systems, policy makers might encourage principals to develop internal accountability systems, which refer to the ability of the school to respond to external pressure in a way that improves its performance. As I identify the results, external accountability systems may have less effect on principals' behaviors. For principals to positively and actively respond to accountability, internal accountability systems may be necessary because they can make a positive impact on teachers' teaching practices because the systems "reflect an alignment within the school of personal responsibility and collective expectations - regardless of the external policy" (Abelmann, Elmore, Even, Kenyon, & Marshall, 1999, p. 38).

However, internal accountability systems do not necessarily develop as the result of the external accountability system (Gonzalez & Firestone, 2013; McGuinn, 2012). In order to create internal accountability, policy makers and school districts should provide sufficient workshops, professional development, and templates for the standards and curriculum of states' accountability policies. Rather than the signal of states' accountability policies, educational resources and school staff capacities are more useful for principals to understand and implement states' accountability policies (B. Berry et al., 2003).

In addition, principals might focus on internalizing the external expectations for the school and share responsibility with their staffs to emphasize students' outcomes (Knapp & Feldman, 2012). Developing new teachers, sustaining instructional success, implementing

curricular innovations, and changing the school-community relationship can be effective methods to enhance professional responsibility of relationship (Jacobson, Johnson, Ylimaki, & Giles, 2009; Polk, 2006).

Second, policy makers can encourage professional development programs for principals. Principals' experiences participating in professional development programs may be an effective method for principals to respond actively and positively to accountability policies. Professional development programs provide not only a better understanding of content and instruction (O'Donnell & White, 2005), but they also offer advocacy and outreach to professional organizations for school principals (Keith, 2011). Principals can improve their abilities for making and evaluating decisions adhering to states' accountability standards through professional development related to data management and analysis (Adamowki, Therriault, & Cavanna, 2007).

Third, it is necessary for principals to emphasize long-term goals. As the results of this study, proficiency performance standards can provide positive effects, although AMO strength and high school graduation exit exams provided negative effects. Long-term goals can be motivations, compared to short-term goals, which are considered as pressure. Therefore, principals with long-term points of view implement educational activities that have high yield.

Fourth, in order to enhance teacher autonomy, principals must invest time and effort to instruction. Principals' reported influence on instruction was positively related to teacher curriculum autonomy and instructional autonomy. Therefore, principals need to develop necessary knowledge and skills to act goals for meet curriculum standards and to evaluate teachers.

Fifth, principals might want to match scheduling of professional days to the focus on the

developmental programs. As confirmed in the results, when principals provide professional days before the school year, teachers can participate in professional development programs related to content. Because the provision of professional days before the school year can be an effective method for teachers to focus on professional development time, principals should implement professional days before the school year, not during the school year.

Lastly, principals need to implement suitable policies for their school contexts. As we observed, there were different influential factors for teachers' work types. If principals would like to improve teacher instructional autonomy and to increase teachers' participation time in classroom management programs, principals should establish healthy school climate, while which may not an effective method for high teacher curriculum autonomy and high participation in professional development content and instruction.

### **3. Conclusion**

Accountability policies have been placed at the most important point since the 1990s. Based on the accountability policies, each state implements its own accountability policies. They established academic content and performance standards, implemented test for all the students in grades 3 through 8 annually, and set up annual measurable objectives in reading and mathematics for districts, schools, and designated student subgroups within schools. The combination of states' decisions on accountability policies, such as performance standards, high school graduation exit exams, and the difference of between starting points and intermediate goals, may lead to the varying strength of the accountability systems in different states. Existing studies have found that the strength of states' accountability systems can affect students' academic outcomes and teachers' instruction.

Based on this study, there are negative effects of states' accountability systems on principals' responses although the effects were not strong. Principals in states with strong accountability systems may have low influence on instruction, and they provide less professional days before and during the school year. Because strong states' accountability systems are like to emphasize high students' test scores, principals in these states tend to focus less on teacher learning.

This study also found that states' accountability systems make effects on a specific domain such as content and curriculum, not instruction. The effects of states' accountability systems are also directly opposed effects of states' accountability policies on teachers' response: long-term goals show positive effects although short-term goals show negative effects. In addition, in schools with specific features, the effects of states' accountability systems are remarkably appeared.

The main goal of accountability policies is to increase students' academic outcomes. Under the accountability systems, principals and teachers implement various activities and behaviors in order to accomplish this goal. However, this study shows that the responses of principals and teachers to strong states' accountability systems might be negative for school staffs and school organization, which can produce low students' academic outcomes. Therefore, it is necessary to modify and develop states' accountability systems in order to create school context that not only students can produce high academic outcomes, but also principals and teachers positively perceive and respond to them.

## APPENDICES

**Appendix A. Proficiency Performance Standards in Fourth and Eighth Grade in Reading**  
**Table VI-1 Proficiency Performance Standards in Fourth and Eighth Grade in Reading**

	Forth grade Reading Performance Standards				Eighth grade Reading Performance Standards					Forth grade Reading Performance Standards				Eighth grade Reading Performance Standards			
	2003	2005	2007	2009	2003	2005	2007	2009		2003	2005	2007	2009	2003	2005	2007	2009
AL			205	207			234	234	MT	229		234	235	253		250	246
AK	223	222	216	218	241	230	233	231	NE								
AZ			213	212	256	244	245	241	NV	228	230	224	225			247	246
AK	223	236	229	216	267	254	249	241	NH			239	237			258	256
CA	231	231	226	220	271	262	261	259	NJ		221	220	231	249	250	252	244
CO		201	201	202	229	229	230	228	NM		233	233	236		251	248	246
CT	217	221	220	214	239	242	245	243	NY	213	207	219	207	272	268	260	247
DE			225	220	249	242	240	236	NC	203	203	231	220	226	217	217	246
FL	231	230	230	225	263	265	262	262	ND	234	224	226	225	255	255	251	253
GA	212	215	213	218	230	224	215	209	OH	232	233	225	219		241	240	251
HI		247	238	239	264	262	245	241	OK		218	213	228	238	244	232	249
ID	217	207	217	213	247	235	233	218	OR			220	214	258	254	251	250
IL			208	207	256	245	236	234	PA			223	218	256	258	245	245
IN		225	228	229	257	249	251	255	RI			236	231			253	252
IA	220	219	220	221	253	250	252	248	SC	248	246	245	215	285	276	281	245
KS	226	218	219	217	253	242	241	236	SD			224	224			249	254
KY			229	223			251	253	TN		200	198	195		222	211	211
LA	221	223	223	221	253	251	246	243	TX	207	219	217	214	221	225	222	201
ME			236	234	274		261	253	UT				225				235
MD		215	206	208	252	245	250	237	VT			239	236			263	259
MA	251	255	254	255			252	249	VA			219	213			239	229
MI	226	222	204	200			238	236	WA	236	236	240	243			253	253
MN			237	233			265	259	WV		215	217	225		228	229	249
MS	205	206	204	223	250	247	251	254	WI	223	225	222	219	232	229	231	232
MO	244	242	245	246			272	267	WY	250	251	216	226	277	278	247	259

**Appendix B. Proficiency Performance Standards in Fourth and Eighth Grade in Mathematics**  
**Table VI-2 Proficiency Performance Standards in Fourth and Eighth Grade in Mathematics**

	Forth grade Math Performance Standards				Eighth grade Math Performance Standards					Forth grade Math Performance Standards				Eighth grade Math Performance Standards			
	2003	2005	2007	2009	2003	2005	2007	2009		2003	2005	2007	2009	2003	2005	2007	2009
AL			205	207			253	246	MT	229		234	235	271		281	285
AK	223	222	216	218	268	268	265	268	NE								
AZ			213	212	300		268	266	NV	228	230	224	225			267	269
AK	223	236	229	216	296	288	277	267	NH			239	237			282	281
CA	231	231	226	220					NJ		221	220	231	278	273	272	272
CO		201	201	202	268	258	259	256	NM		233	233	236		287	285	277
CT	217	221	220	214	258	257	252	251	NY	213	207	219	207	279	275	273	249
DE			225	220	250	252		258	NC	203	203	231	220	247	247	270	253
FL	231	230	230	225	269	269	266	266	ND	234	224	226	225	293	277	279	278
GA	212	215	213	218	255	255	243	247	OH	232	233	225	219		274	265	265
HI		247	238	239	299	296	294	286	OK		218	213	228	256	258	249	269
ID	217	207	217	213	280	266	265	261	OR			220	214	275	269	262	266
IL			208	207	276	276	251	251	PA			223	218	279	272	271	272
IN		225	228	229	269	266	266	273	RI			236	231			279	275
IA	220	219	220	221	266	262	264	263	SC	248	246	245	215	306	305	312	270
KS	226	218	219	217			270	265	SD			224	224			271	271
KY			229	223	291	285	279	273	TN		200	198	195		230	234	229
LA	221	223	223	221	265	264	267	263	TX	207	219	217	214	260	273	268	254
ME			236	234	311		286	284	UT				225				275
MD		215	206	208	286	276	278	271	VT			239	236			284	282
MA	251	255	254	255	299	301	302	300	VA			219	213			259	251
MI	226	222	204	200	278	269	260	253	WA	236	236	240	243			286	288
MN			237	233			286	287	WV		215	217	225		253	253	270
MS	205	206	204	223	261	262	262	264	WI	223	225	222	219	261	263	262	262
MO	244	242	245	246	314	311	289	287	WY	250	251	216	226	297	293	279	278



## Appendix C. Starting Points of 50 States in 2002

**Table VI-3 Starting Points of 50 States in 2002**

	Forth Reading	Forth Math	Eighth Reading	Eighth Math		Forth Reading	Forth Math	Eighth Reading	Eighth Math
AL	68.00	61.00	43.00	48.00	MT	74.00	51.00	74.00	51.00
AK	64.03	54.86	64.03	54.86	NE	62.00	65.00	61.00	58.00
AZ	45.00	50.00	31.00	7.00	NV	30.00	36.00	37.00	32.00
AR	42.40	40.00	35.20	29.10	NH	82.00	76.00	82.00	76.00
CA	13.60	16.00	13.60	16.00	NJ	68.00	53.00	58.00	39.00
CO	76.92	75.86	73.61	59.51	NM	40.85	24.13	36.79	15.28
CT	57.00	65.00	57.00	65.00	NY	122.00	86.00	122.00	86.00
DE	57.00	33.00	57.00	33.00	NC	68.90	65.80	68.90	65.80
FL	31.00	38.00	31.00	38.00	ND	65.10	45.70	61.40	33.30
GA	60.00	50.00	60.00	50.00	OH	62.00	35.90	68.60	37.00
HI	30.00	10.00	30.00	10.00	OK	622.00	648.00	622.00	648.00
ID	66.00	51.00	66.00	51.00	OR	40.00	39.00	40.00	39.00
IL	40.00	40.00	40.00	40.00	PA	45.00	35.00	45.00	35.00
IN	58.80	57.10	58.80	57.10	RI	76.10	61.70	68.00	46.10
IA	64.00	62.00	60.00	58.00	SC	17.60	15.50	17.60	15.50
KS	51.20	46.80	51.20	46.80	SD	65.00	45.00	65.00	45.00
KY	47.27	22.45	45.60	16.49	TN	77.00	72.00	77.00	72.00
LA	36.90	30.10	36.90	30.10	TX	46.80	33.40	46.80	33.40
ME	34.00	12.00	35.00	13.00	UT	65.00	57.00	65.00	57.00
MD	43.80	41.40	43.00	19.00	VT	403.00	390.00	403.00	390.00
MA	70.70	53.00	70.70	53.00	VA	60.70	58.40	60.70	58.40
MI	38.00	47.00	31.00	31.00	WA	52.20	29.70	30.10	17.30
MN	69.50	69.60	64.00	58.30	WV	72.00	67.00	75.00	64.00
MS	66.00	49.00	30.00	27.00	WI	61.00	37.00	61.00	37.00
MO	18.40	8.30	18.40	8.30	WY	30.40	23.80	34.50	25.30

## Appendix D Intermediate Goals of 50 States in 2007

**Table VI-4 Intermediate Goals of 50 States in 2007**

	Forth Reading	Forth Math	Eighth Reading	Eighth Math		Forth Reading	Forth Math	Eighth Reading	Eighth Math
AL	77.00	72.00	59.00	55.00	MT	83.00	68.00	83.00	68.00
AK	77.18	66.09	77.18	66.09	NE	81.00	83.00	81.00	79.00
AZ	56.00	63.20	54.00	38.00	NV	51.70	54.60	51.70	54.60
AR	56.80	55.00	51.40	46.83	NH	86.00	82.00	86.00	82.00
CA	35.20	37.00	35.20	37.00	NJ	82.00	73.00	76.00	62.00
CO	88.46	89.09	86.81	79.75	NM	59.00	44.00	56.00	38.00
CT	79.00	82.00	79.00	82.00	NY	133.00	102.00	133.00	102.00
DE	68.00	50.00	68.00	50.00	NC	76.70	77.20	76.70	77.20
FL	58.00	62.00	58.00	62.00	ND	82.60	72.90	80.70	66.70
GA	73.30	59.50	73.30	59.50	OH	74.60	73.70	79.00	58.00
HI	58.00	46.00	58.00	46.00	OK	914.00	932.00	914.00	932.00
ID	78.00	70.00	78.00	70.00	OR	60.00	59.00	60.00	59.00
IL	62.50	62.50	62.50	62.50	PA	56.00	63.00	56.00	63.00
IN	72.60	71.50	72.60	71.50	RI	84.10	74.50	78.60	64.10
IA	76.00	74.70	73.30	72.00	SC	58.80	57.80	58.80	57.80
KS	75.60	73.40	75.60	73.40	SD	82.00	72.00	82.00	72.00
KY	60.45	41.84	59.20	37.37	TN	89.00	79.00	89.00	79.00
LA	57.90	53.50	57.90	53.50	TX	60.00	50.00	60.00	50.00
ME	50.00	40.00	50.00	40.00	UT	77.00	71.00	77.00	71.00
MD	71.80	69.10	71.10	57.20	VT	435.00	427.00	435.00	427.00
MA	85.40	76.50	85.40	76.50	VA	77.00	75.00	77.00	75.00
MI	59.00	65.00	54.00	54.00	WA	76.10	64.90	65.10	58.70
MN	73.80	73.90	69.20	64.30	WV	76.67	72.50	79.17	70.00
MS	83.00	75.00	65.00	64.00	WI	74.00	58.00	74.00	58.00
MO	51.00	45.00	51.00	45.00	WY	53.60	49.20	56.30	50.20

## Appendix E. Number of High School Teachers Among 50 States

**Table VI-5 Number of High School Teachers Among 50 States**

	The number of high school teachers																Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
AL	2	4	1	2	6	11	10	6	8	7	4	2	1	0	0	0	64
AK	1	2	4	4	6	7	2	3	3	0	2	0	1	0	0	0	35
AZ	3	2	3	8	5	13	8	10	8	6	3	1	1	1	1	0	73
AK	2	7	7	8	6	9	12	6	6	3	1	2	0	0	0	0	69
CA	8	10	8	5	7	13	12	10	6	6	4	3	3	0	0	0	95
CO	2	1	2	8	9	4	4	5	4	3	4	1	0	0	0	0	47
CT	0	0	1	2	6	9	7	4	12	3	1	3	0	0	0	0	48
DE	3	2	1	2	3	8	1	2	2	1	1	0	0	0	0	0	26
DC	16	7	6	1	2	3	2	5	2	5	4	3	7	1	1	0	65
FL	3	0	2	1	5	8	11	9	11	3	1	1	0	1	0	0	56
GA	2	2	2	2	2	4	5	2	2	0	0	0	0	0	0	0	23
HI	1	7	6	2	10	6	9	13	7	1	3	3	0	1	1	0	70
ID	3	3	5	2	11	11	6	4	3	4	4	2	1	1	1	0	61
IL	1	1	2	8	8	7	11	5	11	5	4	1	3	0	0	0	67
IN	1	0	5	3	6	8	4	8	6	3	2	2	1	0	0	0	49
IA	0	3	3	5	4	8	12	6	6	4	3	1	2	3	0	0	60
KS	1	4	4	3	3	10	12	9	3	5	1	2	1	0	0	0	58
KY	1	3	2	7	9	9	3	5	6	1	2	1	1	0	0	0	50
LA	0	4	2	5	6	12	8	8	7	2	3	1	0	0	0	0	58
ME	1	4	7	2	6	10	4	5	5	1	0	0	0	0	0	0	45
MD	0	3	4	3	6	9	7	9	6	4	3	0	1	0	0	0	55
MA	4	6	2	5	8	12	6	4	3	3	1	2	0	0	0	0	56
MI	1	4	3	2	10	3	13	14	6	3	1	4	1	0	1	0	66
MN	2	0	7	4	4	7	8	7	7	3	5	4	1	0	1	0	60
MS	0	2	2	6	12	8	5	7	6	7	8	1	2	2	0	0	68

Table VI-5 (cont'd)

	The number of high school teachers																Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
MO	1	0	4	3	8	3	7	10	1	3	1	4	2	2	1	0	50
MT	0	1	0	3	9	3	9	3	5	1	1	5	1	0	0	1	42
NE	5	2	5	5	10	2	1	5	2	5	3	2	0	0	0	0	47
NV	2	0	2	2	6	8	2	5	6	3	3	0	0	0	0	0	39
NH	5	6	1	2	5	7	6	10	8	1	1	1	0	0	0	0	53
NJ	2	6	6	10	15	5	4	2	1	5	3	1	1	1	1	0	63
NM	1	2	1	5	6	7	3	4	5	8	2	0	1	1	0	0	46
NY	3	4	5	5	9	7	7	10	1	1	0	0	0	0	0	0	52
NC	0	0	2	6	6	2	3	1	2	2	4	1	1	2	1	0	33
ND	2	1	5	2	4	8	10	10	3	2	3	2	3	1	0	0	56
OH	1	7	4	5	6	11	12	19	15	7	5	5	1	0	0	0	98
OK	4	0	4	6	6	8	3	4	4	3	3	1	2	2	0	0	50
OR	0	2	1	1	4	11	12	9	5	7	2	1	1	2	0	0	58
PA	2	4	1	2	7	4	0	1	0	0	0	0	0	0	0	0	21
RI	1	1	1	4	14	6	7	13	8	4	3	2	1	0	0	0	65
SC	1	0	2	3	2	4	5	5	2	0	1	1	2	0	1	0	29
SD	0	1	1	6	7	6	10	5	6	5	3	2	1	0	0	0	53
TN	6	2	7	12	12	9	4	6	9	3	4	3	0	2	0	0	79
TX	0	2	7	7	7	6	11	10	2	1	1	1	0	0	0	0	55
UT	1	4	5	0	5	6	4	2	2	0	0	1	0	0	0	0	30
VT	1	1	5	4	7	11	8	8	4	1	0	0	0	0	0	0	50
VA	3	3	4	6	8	4	8	10	5	3	2	2	0	0	0	0	58
WA	6	2	5	5	6	8	5	6	10	3	2	6	4	0	0	0	68
WV	1	3	1	6	7	10	13	7	4	4	3	2	1	0	0	0	62
WI	0	2	3	3	4	6	7	3	3	2	1	1	0	0	0	0	35
WY	1	2	2	4	8	3	7	4	2	2	1	0	0	0	0		36
	106	137	173	213	340	371	343	334	259	157	116	84	49	23	10	1	2716

## Appendix F. Principals Responses by States

**Table VI-6 Principals Responses by States**

	Influence on instruction	Facilitating teacher learning	Provision of professional days		State	Influence on instruction	Facilitating teacher learning	Provision of professional days	
			Before the school year	During the school year				Before the school year	During the school year
AL	3.578	2.317	0.984	0.889	MT	3.667	2.587	1.000	1.000
AK	3.477	2.171	1.000	0.943	NE	3.729	2.558	0.977	0.953
AZ	3.615	2.573	0.960	0.733	NV	3.619	2.239	0.978	0.804
AK	3.600	1.924	1.000	0.864	NH	3.699	2.500	0.947	0.947
CA	3.657	2.950	0.921	0.723	NJ	3.576	2.356	0.864	0.915
CO	3.528	2.620	0.960	0.980	NM	3.691	2.603	0.985	0.956
CT	3.692	2.596	1.000	0.962	NY	3.767	2.519	0.885	0.923
DE	3.667	2.269	1.000	1.000	NC	3.574	2.250	0.942	0.865
FL	3.652	2.552	0.970	0.910	ND	3.745	2.394	1.000	0.939
GA	3.672	2.414	0.983	0.966	OH	3.548	2.077	0.846	0.904
HI	3.692	2.731	0.962	1.000	OK	3.714	1.947	0.989	0.915
ID	3.662	2.364	0.939	0.985	OR	3.569	2.367	0.980	0.959
IL	3.801	2.657	0.970	0.896	PA	3.693	2.197	1.000	1.000
IN	3.662	2.354	0.800	0.892	RI	3.587	2.087	0.783	0.739
IA	3.653	2.449	1.000	1.000	SC	3.736	2.224	0.985	0.970
KS	3.607	2.407	0.983	0.983	SD	3.770	2.036	1.000	0.929
KY	3.640	1.880	1.000	0.920	TN	3.727	1.981	0.962	0.885
LA	3.607	2.418	0.982	0.909	TX	3.679	2.667	0.988	0.951
ME	3.640	2.918	0.984	0.918	UT	3.626	2.286	0.982	0.946
MD	3.518	2.511	0.978	0.933	VT	3.731	2.355	0.968	0.968
MA	3.784	2.196	0.893	0.857	VA	3.724	2.569	0.980	0.961
MI	3.519	1.931	0.948	0.948	WA	3.576	2.678	1.000	0.949
MN	3.637	2.412	0.985	0.971	WV	3.662	2.087	1.000	0.971
MS	3.656	2.362	1.000	0.862	WI	3.621	2.469	0.969	0.844
MO	3.633	2.551	0.957	0.942	WY	3.596	2.368	1.000	0.974

## Appendix G. Teachers' Responses by States

**Table VI-7 Teachers' Responses by States**

	Autonomy		Professional development time				Autonomy		Professional development time		
	Curriculum	Instruction	Content	Instruction	Classroom management		Curriculum	Instruction	Content	Instruction	Classroom management
AL	2.594	3.644	1.910	0.678	0.723	MT	3.213	3.759	2.311	0.626	0.695
AK	3.151	3.731	2.294	0.538	0.521	NE	3.253	3.729	2.085	0.745	0.610
AZ	2.917	3.703	2.140	0.941	0.549	NV	2.785	3.596	2.204	0.707	0.780
AK	2.939	3.689	2.651	1.103	0.963	NH	3.091	3.659	2.470	0.580	0.475
CA	2.781	3.670	2.137	0.813	0.631	NJ	2.874	3.604	1.694	0.351	0.570
CO	2.967	3.675	2.412	1.146	0.407	NM	3.087	3.674	1.737	0.920	0.357
CT	3.054	3.652	1.715	0.599	0.310	NY	3.067	3.627	1.989	0.504	0.466
DE	2.758	3.552	1.871	0.710	0.790	NC	2.721	3.625	1.755	0.723	0.665
FL	2.708	3.695	2.275	1.668	0.646	ND	3.372	3.796	2.094	0.661	0.531
GA	2.535	3.606	1.906	0.437	0.549	OH	3.199	3.706	1.813	0.557	0.545
HI	2.883	3.597	1.662	0.870	0.519	OK	3.104	3.707	1.958	0.465	0.682
ID	3.037	3.748	2.180	0.587	0.654	OR	3.134	3.741	2.134	1.290	0.601
IL	3.170	3.719	1.826	0.830	0.681	PA	3.153	3.695	1.877	0.877	0.574
IN	3.118	3.676	1.545	0.696	0.491	SC	2.751	3.628	1.910	0.836	0.630
IA	3.358	3.736	1.933	1.253	0.494	SD	3.247	3.713	2.116	0.902	0.768
KS	3.066	3.688	2.050	0.991	0.737	TN	2.801	3.720	1.954	0.518	0.803
KY	2.928	3.587	2.147	0.938	0.598	TX	2.663	3.550	2.484	0.641	0.884
LA	2.626	3.618	1.936	0.717	0.775	UT	3.109	3.823	2.496	0.750	0.719
ME	3.314	3.704	2.114	0.978	0.298	VT	3.304	3.709	2.554	0.793	0.543
MD	2.299	3.563	2.091	0.753	0.578	VA	2.642	3.589	2.187	0.651	0.572
MA	3.104	3.667	2.352	0.578	0.522	WA	2.969	3.670	2.217	0.972	0.510
MI	2.882	3.659	2.041	0.774	0.421	WV	2.770	3.697	1.811	0.468	0.510
MN	3.290	3.755	2.068	1.047	0.660	WI	3.241	3.697	2.006	0.877	0.464
MS	2.828	3.628	1.626	0.554	0.742	WY	3.194	3.647	1.926	1.123	0.660
MO	3.022	3.645	2.136	0.726	0.770						

## BIBLIOGRAPHY

## BIBLIOGRAPHY

- Abelmann, C., Elmore, R. F., Even, J., Kenyon, S., & Marshall, J. (1999). *When Accountability Knocks, Will Anyone Answer?* Philadelphia, PA: CPRE Publications, University of Pennsylvania, Graduate School of Education.
- Abrams, L. M., Pedulla, J. J., & Madaus, G. F. (2003). Views from the Classroom: Teachers' Opinions of Statewide Testing Programs. *Theory Into Practice*, 42(1), 18-29.
- Adamowki, S., Therriault, S. B., & Cavanna, A. P. (2007). *The Autonomy Gap: Barriers to Effective School Leadership*. Washington, DC: Thomas B. Fordham Foundation & Institute.
- Amrein, A. L., & Berliner, D. C. (2002). *The impact of high-stakes tests on student academic performance: An analysis of NAEP results in states with high-stakes tests and ACT, SAT, and AP test results in states with high school graduation exams*. Tempe, AZ: Education Policy Studies Laboratory, Arizona State University.
- Anagnostopoulos, D. (2006). "Real Students" and "True Demotes": Ending Social Promotion and the Moral Ordering of Urban High Schools. *American Educational Research Journal*, 43(1), 5-42.
- Arbogast, A. D. (2004). *Supporting professional learning in an era of accountability: The elementary school principal perspective*. Unpublished doctoral dissertation, University of Maryland, College Park, MD.
- Arnold, M. L., Newman, J. H., Gaddy, B. B., & Dean, C. B. (2005). A Look at the Condition of Rural Education Research: Setting a Direction for Future Research. *Journal of Research in Rural Education*, 20(6), 1-25.
- Assor, A., & Oplatka, I. (2003). Towards a comprehensive conceptual framework for understanding principals' personal-professional growth. *Journal of Educational Administration*, 41(5), 471-497.
- Au, W. (2007). High-Stakes Testing and Curricular Control: A Qualitative Metasynthesis. *Educational Researcher*, 36(5), 258-267.
- Ayers, J. (2011). *No Child Left Behind Waiver Applications: Are They Ambitious and Achievable?* Washington, DC: Center for American Progress.



- Bae, S. (2008). District Capacity and Accountability: Professional Development as Reform Tool. *Research in the Sociology of Education*, 16, 189-207.
- Bakkenes, I., Vermunt, J. D., & Wubbels, T. (2010). Teacher learning in the context of educational innovation: Learning activities and learning outcomes of experienced teachers. *Learning and Instruction*, 20(6), 533-548.
- Ball, D. L., & Cohen, D. K. (1999). Developing Practice, Developing Practitioners: Toward a Practice-Based Theory of Professional Education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the Learning Profession: Handbook of Policy and Practice* (pp. 3–32). San Francisco, CA: San Francisco: Jossey-Bass.
- Ballou, D., & Springer, M. G. (2009). *Achievement Trade-Offs and No Child Left Behind*. Washington, DC: Urban Institute.
- Bandeira de Mello, V., Blankenship, C., & McLaughlin, D. (2009). *Mapping State Proficiency Standards Onto NAEP Scales: 2005-2007*. Jessup, MD: National Center for Education Statistics.
- Bell, B. A., Ferron, J. M., & Kromrey, J. D. (2008). *Cluster size in multilevel models: The impact of sparse data structures on point and interval estimates in two-level models*. Paper presented at the JSM Proceedings.
- Bell, B. A., Morgan, G. B., Kromrey, J. D., & Ferron, J. M. (2010). *The Impact of Small Cluster Size on Multilevel Models: A Monte Carlo Examination of Two-Level Models with Binary and Continuous Predictors*. Paper presented at the JSM Proceedings.
- Berry, B., Turchi, L., Johnson, D., Hare, D., Owens, D. D., & Clements, S. (2003). *The Impact of High Stakes Accountability on Teachers' Professional Development: Evidence from the South*. Chapel Hill, NC.: Southeast Center for Teaching Quality.
- Berry, K. S., & Herrington, C. D. (2011). States and their struggles with NCLB: Does the Obama Blueprint get it right? *Peabody Journal of Education*, 86(3), 272-290.
- Bishop, J. H., & Mane, F. (2001). The impacts of minimum competency exam graduation requirements on high school graduation, college attendance and early labor market success. *Labour Economics*, 8(2), 203-222.
- Bishop, J. H., Moriarty, J. Y., & Mane, F. (2000). Diplomas for Learning, Not Seat Time: The

- Impacts of New York Regents Examinations. *Economics of Education Review*, 19(4), 333-349.
- Bloom, C. M., & Owens, E. W. (2013). Principals' Perception of Influence on Factors Affecting Student Achievement in Low- and High-Achieving Urban High Schools. *Education and Urban Society*, 45(2), 208-233.
- Booher-Jennings, J. (2005). Below the Bubble: "Educational Triage" and the Texas Accountability System. *American Educational Research Journal*, 42(2), 231-268.
- Boser, U. (2001). Pressure without support. *Education Week*, 20(17), 68-71.
- Bottoms, G. (2003). What School Principals Need To Know about Curriculum and Instruction. *ERS Spectrum*, 21(1), 29-31.
- Brown, D. W., Carr, R. E., Perry, C. M., & McIntire, W. G. (1996). Principals' Perceptions of Community and Staff Involvement in Shared Decision Making. *Journal of Research in Rural Education*, 12(1), 17-24.
- Brunetti, G. J. (2001). Why Do They Teach? A Study of Job Satisfaction among Long-Term High School Teachers. *Teacher Education Quarterly*, 28(3), 49-74.
- Bubb, S., & Earley, P. (2013). The use of training days: finding time for teachers' professional development. *Educational Research*, 55(3), 236-248.
- Buchholz, C., & List, K. L. (2009). A Place for Learning. *Principal Leadership*, 9(7), 38-42.
- Byrne, J. L. (2009). *Elementary teachers' perceptions of autonomy in light of the standards movement and No Child Left Behind*. 3387249, University of Wyoming, United States -- Wyoming.
- Carnoy, M. (2005). Have State Accountability and High-Stakes Tests Influenced Student Progression Rates in High School? *Educational Measurement, Issues and Practice*, 24(4), 19-31.
- Carnoy, M., & Loeb, S. (2002). Does External Accountability Affect Student Outcomes? A Cross-State Analysis. *Educational Evaluation and Policy Analysis*, 24(4), 305-331.

- Chase, F. S. (1971). Problems of autonomy and accountability in government contracts for research and development in education. *The Dilemma of accountability in modern government; independence versus control*, 103-117.
- Chudowsky, N., Kober, N., Gayler, K. S., & Hamilton, M. (2002). *State High School Exit Exams: A Baseline Report*.
- Clarke, M., Shore, A., Rhoades, K., Abrams, L., Miao, J., & Li, J. (2003). *Perceived Effects of State-Mandated Testing Programs on Teaching and Learning: Findings from Interviews with Educators in Low-, Medium-, and High-Stakes States*. Chestnut Hill, MA: National Board on Educational Testing and Public Policy.
- Clarke, P., & Wheaton, B. (2007). Addressing data sparseness in contextual population research: using cluster analysis to create synthetic neighborhoods. *Sociological Methods and Research*, 35(3), 311-351.
- Clotfelter, C. T., & Ladd, H. F. (1996). Recognizing and Rewarding Success in Public Schools. In H. F. Ladd (Ed.), *Holding schools accountable : performance-based reform in education* (pp. 23-64). Washington, D.C.: Brookings Institution.
- Cocke, E. F., Buckley, J., & Scott, M. A. (2011). *Accountability and Teacher Practice: Investigating the Impact of a New State Test and the Timing of State Test Adoption on Teacher Time Use*. Paper presented at the 2011 SREE Conference, Washington, D.C.
- Coleman, J. S. (1966). *Equality of educational opportunity*: United States Government Printing Office.
- Corcoran, S. P., Schwartz, A. E., & Weinstein, M. (2012). Training Your Own: The Impact of New York City's Aspiring Principals Program on Student Achievement. *Educational Evaluation and Policy Analysis*, 34(2), 232-253.
- Cox, J. H., & Witko, C. (2011). *Accountability and the "Narrowing" of the Curriculum*. Paper presented at the The State politics and policy conference, Hanover, NH.
- Crocco, M. S., & Costigan, A. T. (2007). The narrowing of curriculum and pedagogy in the age of accountability: Urban educators speak out. *Urban Education*, 42(6), 512-535.
- Croft, A., Coggs, J. G., Dolan, M., & Powers, E. (2010). *Job-Embedded Professional Development: What It Is, Who Is Responsible, and How to Get It Done Well. Issue Brief*.

National Comprehensive Center for Teacher Quality. 1000 Thomas Jefferson Street NW, Washington, DC 20007.

Crowe, E. (2011). *Race to the Top and Teacher Preparation: Analyzing State Strategies for Ensuring Real Accountability and Fostering Program Innovation*.

Daniels, D. M. (2009). *Leadership, Learning and School Change: The Elementary Principal's Role in Teacher Professional Development*. Capella University, Minneapolis, MN.

Davidson, E., Reback, R., Rockoff, J. E., & Schwartz, H. L. (2013). *Fifty Ways to Leave a Child Behind: Idiosyncrasies and Discrepancies in States' Implementation of NCLB*. National Bureau of Economic Research.

Davison, M. L., Kwak, N., Seo, Y. S., & Choi, J. (2002). Using hierarchical linear models to examine moderator effects: Person-by-organization interactions. *Organizational Research Methods*, 5(3), 231-254.

Dean, C. B. (2001). *State Policy Support for Professional Development in the Central Region*. Aurora, CO: Mid-Continent Research for Education and Learning.

Dee, T. S. (2002). *Standards and Student Outcomes: Lessons from the "First Wave" of Education Reform*. Paper presented at the "Taking Account of Accountability: Assessing Politics and Policy," Cambridge, MA.

Dee, T. S., & Jacob, B. A. (2011). The impact of no Child Left Behind on student achievement. *Journal of Policy Analysis and Management*, 30(3), 418-446.

Dee, T. S., Jacob, B. A., Hoxby, C. M., & Ladd, H. F. (2010). The Impact of No Child Left Behind on Students, Teachers, and Schools. *Brookings Papers on Economic Activity*, 149-207.

DeMoss, K. (2002). Leadership styles and high-stakes testing: Principals make a difference. *Education and Urban Society*, 35(1), 111-132.

Desimone, L. M. (2013). Teacher and Administrator Responses to Standards-Based Reform. *Teachers College Record*, 115(8), 1.

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of

- Professional Development on Teachers' Instruction: Results from a Three-year Longitudinal Study. *Educational Evaluation and Policy Analysis*, 24(2), 81-112.
- Desimone, L. M., Smith, T. M., & Phillips, K. J. R. (2007). Does Policy Influence Mathematics and Science Teacher's Participation in Professional Development? *Teachers College Record*, 109(5), 1086-1122.
- Diamond, J. B. (2007). Where the Rubber Meets the Road: Rethinking the Connection Between High-Stakes Testing Policy and Classroom Instruction. *Sociology of Education*, 80(4), 285-313.
- Diamond, J. B. (2012). Accountability policy, school organization, and classroom practice: Partial recoupling and educational opportunity. *Education and Urban Society*, 44(2), 151-182.
- Drage, K. (2010). Professional Development: Implications for Illinois Career and Technical Education Teachers. *Journal of Career and Technical Education*, 25(2), 24-37.
- Drago-Severson, E. (2012). New Opportunities for Principal Leadership: Shaping School Climates for Enhanced Teacher Development. *Teachers College Record*, 114(3), 6.
- Eden, D. (2001). Who Controls the Teachers? Overt and Covert Control in Schools. *Educational Management & Administration*, 29(1), 97-111.
- Education, N. C. o. E. i. (1983). *A Nation at Risk*. Washington, DC: GPO.
- Elmore, R. F. (2005). Accountable Leadership. *The Educational Forum*, 69(2), 134-142.
- Elmore, R. F., Ableman, C., & Fuhrman, S. H. (1996). The new accountability in state education reform: From process to performance. In H. F. Ladd (Ed.), *Holding schools accountable : performance-based reform in education* (pp. 65-98). Washington, D.C.: Brookings Institution.
- Englert, K., Fries, D., Martin-Glenn, M., & Douglas, B. (2007). How are Educators Using Data?: A Comparative Analysis of Superintendent, Principal, and Teachers' Perceptions of Accountability Systems. *International Journal of Educational Policy & Leadership*, 2(4), 1-12.

- Erpelding, C. J. (1999). *School vision, teacher autonomy, school climate, and student achievement in elementary schools*. Ed.D., University of Northern Iowa, Ann Arbor.
- Erpenbach, W. J. (2011). *Statewide Educational Accountability Systems Under the NCLB ACT--A Report on 2009 and 2010 Amendments to State Plans*. Washington, DC: Council of Chief State School Officers.
- Erpenbach, W. J., Forte-Fast, E., & Potts, A. (2003). *Statewide Educational Accountability under NCLB. Central Issues Arising from An Examination of State Accountability Workbooks and U.S. Department of Education Reviews under the No Child Left Behind Act of 2001*. Washington, DC: Council of Chief State School Officers.
- Fast, E. F., & Erpenbach, W. J. (2004). *Revisiting Statewide Educational Accountability Under NCLB: A Summary of State Requests in 2003-2004 for Amendments to State Accountability Plans*. Washington, DC: Council of Chief State School Officers.
- Feng, L., Figlio, D. N., & Sass, T. R. (2010). *School Accountability and Teacher Mobility*. Washington, DC: National Center for Analysis of Longitudinal Data in Education Research. The Urban Institute.
- Figlio, D. N., Rouse, C. E., & Schlosser, A. (2009). *Leaving No Child Behind: Two Paths to School Accountability*. Washington, DC: Urban Institute.
- Finn, C. E., Jr. (2012, March 1). The War Against the Common Core. Retrieved from <http://www.edexcellence.net/commentary/education-gadfly-weekly/2012/march-1/the-war-against-the-common-core-1.html>
- Finn, C. E., Jr., & Kanstoroom, M. (2001). State Academic Standards. *Brookings papers on education policy*, 131-179.
- Finnigan, K. S. (2012). Principal Leadership in Low-Performing Schools: A Closer Look Through the Eyes of Teachers. *Education and Urban Society*, 44(2), 183-202.
- Finnigan, K. S., & Gross, B. (2007). Do Accountability Policy Sanctions Influence Teacher Motivation? Lessons From Chicago's Low-Performing Schools. *American Educational Research Journal*, 44(3), 594-629.
- Firestone, W. A., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational*

*Evaluation and Policy Analysis*, 20(2), 95-113.

- Firestone, W. A., & Shipps, D. (2005). How do leaders interpret conflicting accountabilities to improve students learning? In W. A. Firestone & C. Riehl (Eds.), *A new agenda for research in educational leadership* (pp. 81-100). New York: Teachers College Press.
- Forte, E., & Erpenbach, W. J. (2006). *Statewide Educational Accountability Under the No Child Left Behind Act: A Report on 2006 Amendments to State Plans. A Summary of State Requests in 2005-06 for Amendments to Their Educational Accountability Systems Under NCLB*. Washington, DC Council of Chief State School Officers.
- Foy, L. L. (2008). *Principals' perceptions of high-stakes accountability testing policies' influence on their behavior as instructional leaders*. AAI3300349, Temple University Ambler, PA
- Fuhrman, S. H. (1999). *The New Accountability*. Philadelphia, PA: CPRE Publications.
- Fuhrman, S. H., Clune, W. H., & Elmore, R. F. (1988). Research on Education Reform: Lessons on the Implementation of Policy. *Teachers College Record*, 90(2), 237-257.
- Fuhrman, S. H., & Elmore, R. F. (2004). *Redesigning accountability systems for education*. New York, NY: Teachers College Press.
- Fullan, M. (2001). *The new meaning of educational change* (3rd ed.). New York: Teachers College Press.
- Gamoran, A., & Dreeben, R. (1986). Coupling and Control in Educational Organizations. *Administrative Science Quarterly*, 31(4), 612-632.
- Gardiner, M. E., Canfield-Davis, K., & Anderson, K. L. (2009). Urban School Principals and the 'No Child Left Behind' Act. *The Urban Review*, 41(2), 141-160.
- Garet, M. S., Porter, A. C., Desimone, L. M., Birman, B. F., & Yoon, K. S. (2001). What Makes Professional Development Effective? Results from a National Sample of Teachers. *American Educational Research Journal*, 38(4), 915-945.
- Garvin, N. M. (2007). *Teacher autonomy: Distinguishing perceptions by school cultural characteristics*. Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia, PA.

- Gavin, M. B., & Hofmann, D. A. (2002). Using hierarchical linear modeling to investigate the moderating influence of leadership climate. *The Leadership Quarterly*, 13(1), 15-33.
- Gayler, K. (2005). *How Have High School Exit Exams Changed Our Schools?: Some Perspectives from Virginia & Maryland*. Center on Education Policy.
- Goertz, M. E. (2001). The Federal Role in an Era of Standards-Based Reform *The Future of the Federal Role in Elementary and Secondary Education: A Collection of Papers*. Washington, DC: Center for Education Policy.
- Goertz, M. E., & Duffy, M. C. (2001). *Assessment and Accountability Systems in the 50 States, 1999-2000. CPRE Research Report Series*. Philadelphia, PA: Consortium for Policy Research in Education, Graduate School of Education, University of Pennsylvania.
- Gonzalez, R. A. (2012). *Educational tug-of-war: Principal leadership and accountability*. Unpublished doctoral dissertation, Rutgers University, New Brunswick, NJ.
- Gonzalez, R. A., & Firestone, W. A. (2013). Educational tug-of-war: internal and external accountability of principals in varied contexts. *Journal of Educational Administration*, 51(3), 383-406.
- Graczewski, C., Knudson, J., & Holtzman, D. J. (2009). Instructional Leadership in Practice: What Does It Look Like, and What Influence Does It Have? *Journal of Education for Students Placed at Risk (JESPAR)*, 14(1), 72-96.
- Grissmer, D., Flanagan, A., Kawata, J., & Williamson, S. (2000). *Improving Student Achievement: What State NAEP Test Scores Tell Us*. Santa Monica, CA: RAND.
- Gross, B., & Goertz, M. E. (2005). *Holding High Hopes: How High Schools Respond to State Accountability Policies*. Philadelphia, PA: Consortium for Policy Research in Education. University of Pennsylvania.
- Hallett, T. (2010). The Myth Incarnate: Re-coupling Processes, Turmoil, and Inhabited Institutions in an Urban Elementary School. *American Sociological Review*, 75(1), 52-74.
- Hallinger, P., & Murphy, J. (1986). *Instructional Leadership in Effective Schools*.
- Hamilton, L. S., Stecher, B. M., Marsh, J. A., McCombs, J. S., Robyn, A., Russell, J., . . . Barney,



- H. (2007). *Standards-Based Accountability Under No Child Left Behind: Experiences of Teachers and Administrators in Three States*. Santa Monica, CA: The RAND Corporation.
- Hamilton, L. S., Stecher, B. M., Russell, J. L., Marsh, J. A., & Miles, J. (2008). Accountability and Teaching Practices: School-Level Actions and Teacher Responses. *Research in the Sociology of Education*, 16, 31-66.
- Haney, W. (2000). The Myth of the Texas Miracle in Education. *Education Policy Analysis Archives*, 8(41).
- Hannaway, J., & Hamilton, L. (2008). *Performance-Based Accountability Policies: Implications for School and Classroom Practices*. Washington, DC: The Urban Institute.
- Hanushek, E. A., & Raymond, M. E. (2004). The effect of school accountability systems on the level and distribution of student achievement. *Journal of the European Economic Association*, 2(2-3), 406-415.
- Hanushek, E. A., & Raymond, M. E. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, 24(2), 297-327.
- Harris, D. M. (2012). Varying teacher expectations and standards: Curriculum differentiation in the age of standards-based reform. *Education and Urban Society*, 44(2), 128-150.
- Harris, D. N., & Herrington, C. D. (2006). Accountability, Standards, and the Growing Achievement Gap: Lessons from the Past Half-Century. *American Journal of Education*, 112(2), 209-238.
- Heinecke, W., Curry-Corcoran, D. E., & Moon, T. R. (2003). U. S. Schools and the new standards and accountability initiative. In D. L. Duke (Ed.), *Educational leadership in an age of accountability : the Virginia experience* (pp. 7-35). Albany, NY: State University of New York Press.
- Hemmings, B., & Kay, R. (2010). Prior Achievement, Effort, and Mathematics Attitude as Predictors of Current Achievement. *Australian Educational Researcher*, 37(2), 41-58.
- Henne, M. K., & Jang, H. (2008). Raising Achievement or Closing Gaps? Identifying Effective Accountability Tools. *Research in the Sociology of Education*, 16, 133-155.

- Hill, H. C. (2007). Learning in the Teaching Workforce. *Future of Children*, 17(1), 111-127.
- Hoffman, J. V., Assaf, L. C., & Paris, S. G. (2001). High-Stakes Testing in Reading: Today in Texas, Tomorrow? *Reading Teacher*, 54(5), 482-492.
- Hofmann, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. *Journal of Management*, 24(5), 623-641.
- Hofmann, D. A., Griffin, M. A., & Gavin, M. B. (2000). The application of hierarchical linear modeling to organizational research. In K. J. Klei & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*. (pp. 467-511). San Francisco, CA: Jossey-Bass
- Hollingworth, L., Dude, D. J., & Shepherd, J. K. (2010). Pizza Parties, Pep Rallies, and Practice Tests: Strategies Used by High School Principals to Raise Percent Proficient. *Leadership and Policy in Schools*, 9(4), 462-478.
- Holme, J. J. (2008). High Stakes Diplomas: Organizational Responses to California's High School Exit Exam. *Research in the Sociology of Education*, 16, 157-188.
- Hood, A. P. (2012). *The increasing standardization of curriculum and instruction in two central-Iowa elementary schools and its effect on teacher autonomy and creativity*. Unpublished doctoral dissertation, Drake University, Des Moines, IA.
- Jackson, A. S. (2006). *Teacher perceptions of professional development in Buncombe County schools*. Ed.D., Western Carolina University, Ann Arbor.
- Jacob, B. A. (2001). Getting tough? The impact of high school graduation exams. *Educational Evaluation and Policy Analysis*, 23(2), 99-121.
- Jacob, B. A. (2005). Accountability, Incentives and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools. *Journal of Public Economics*, 89(5-6), 761-796.
- Jacobson, S. L., Johnson, L., Ylimaki, R., & Giles, C. (2009). Sustaining success in an American school: a case for governance change. *Journal of Educational Administration*, 47(6), 753-764.
- Jasper, B., & Taube, S. (2004). Action research of elementary teachers, problem-solving skills

- before and after focused professional development. *Teacher Education and Practice*, 17(3), 299-310.
- Jenkins, J., & Pfeifer, R. S. (2012). The Principal as Curriculum Leader. *Principal Leadership*, 12(5), 30-34.
- Jones, B. D., & Egley, R. J. (2010). Mixed Feelings: Principals React to Test-Based Accountability. *ERS Spectrum*, 28(2), 17-26.
- Joyce, B. R., & Showers, B. (2002). *Student Achievement Through Staff Development* (3rd ed.). Alexandria, VA: Association for Supervision & Curriculum Development.
- Kee, A. N. (2012). Feelings of Preparedness Among Alternately Certified Teachers: What Is the Role of Program Features? *Journal of Teacher Education*, 63(1), 23-38.
- Keith, D. L. (2011). PRINCIPAL DESIRABILITY FOR PROFESSIONAL DEVELOPMENT. *Academy of Educational Leadership Journal*, 15(2), 95-128.
- Kelley, C. (1999). The Motivational Impact of School-Based Performance Awards. *Journal of Personnel Evaluation in Education*, 12(4), 309-326.
- Kelley, C., Kimball, S., & Conley, S. (2000). Payment for Results: Effects of the Kentucky and Maryland Group-based Performance Award Programs. *Peabody Journal of Education*, 75(4), 159-199.
- Kelley, C., & Protsik, J. (1997). Risk and Reward: Perspectives on the Implementation of Kentucky's School-Based Performance Award Program. *Educational Administration Quarterly*, 33(4), 474-505.
- Klein, S., Hamilton, L., McCaffrey, D., & Stecher, B. (2000). What Do Test Scores in Texas Tell Us? *Education Policy Analysis Archives*, 8(49).
- Knapp, M. S., & Feldman, S. B. (2012). Managing the Intersection of Internal and External Accountability: Challenge for Urban School Leadership in the United States. *Journal of Educational Administration*, 50(5), 666-694.
- Knobl, S. J. (2010). *Perceptions of the roles, professional development, challenges, and frustrations of high school principals*. Unpublished doctoral dissertation University of

South Florida, Tampa, FL.

Kober, N., Chudowsky, N., & Chudowsky, V. (2008). *Has Student Achievement Increased since 2002? State Test Score Trends through 2006-07*. Washington, DC: Center on Education Policy.

Kober, N., & Rentner, D. S. (2012). *Year Two of Implementing the Common Core State Standards: States' Progress and Challenges*. Washington, DC: Center on Education Policy.

Koretz, D. M. (2005). Alignment, High Stakes, and the Inflation of Test Scores. *Yearbook of the National Society for the Study of Education*, 104(2), 99-118.

Koretz, D. M., & Training, R. I. o. E. (1996). *Perceived Effects of the Kentucky Instructional Results Information System (KIRIS)*. Santa Monica, CA: The Rand Corporation.

Kose, B. W. (2009). The Principal's Role in Professional Development for Social Justice: An Empirically-Based Transformative Framework. *Urban Education*, 44(6), 628-663.

Kozar, V. C. F. (2011). *Accountability from the inside out: A case study of isolation and autonomy*. Ph.D., University of Pittsburgh, United States -- Pennsylvania.

Kreft, I., & Leeuw, J. d. (1998). *Introducing multilevel modeling*. London ; Thousand Oaks, Calif.: Sage.

Ladd, H. F. (1996). *Holding schools accountable : performance-based reform in education*. Washington, D.C.: Brookings Institution.

Ladd, H. F., & Lauen, D. L. (2010). Status versus growth: the distributional effects of school accountability policies. *Journal of Policy Analysis and Management*, 29(3), 426-450.

Ladd, H. F., & Zelli, A. (2002). School-based accountability in North Carolina: The responses of school principals. *Educational Administration Quarterly*, 38(4), 494-529.

Lam, Y. L. J. (2005). School organizational structures: effects on teacher and student learning. *Journal of Educational Administration*, 43(4/5), 387-401.

Lambert, L. (2003). *Leadership Capacity for Lasting School Improvement*. Alexandria, VA:

Association for Supervision and Curriculum Development (ASCD).

Le Floch, K. C., Martinez, J. F., O'Day, J., Stecher, B. M., Taylor, J., Cook, A., . . . Garet, M. (2007). *State and Local Implementation of the No Child Left Behind Act: Volume III - Accountability Under NCLB: Interim Report*. Santa Monica, CA: The Rand Corporation.

Lee, J. (2006). *Tracking Achievement Gaps and Assessing the Impact of NCLB on the Gaps: An In-Depth Look into National and State Reading and Math Outcome Trends*. Cambridge, MA: Harvard Education Publishing Group.

Lee, J. (2010). Trick or treat: new ecology of education accountability system in the USA. *Journal of Education Policy*, 25(1), 73-93.

Lee, J., & Reeves, T. (2012). Revisiting the impact of NCLB high-stakes school accountability, capacity, and resources: State NAEP 1990 -2009 reading and math achievement gaps and trends. *Educational Evaluation and Policy Analysis*, 34(2), 209-231.

Lee, J., & Wong, K. K. (2004). The Impact of Accountability on Racial and Socioeconomic Equity: Considering Both School Resources and Achievement Outcomes. *American Educational Research Journal*, 41(4), 797-832.

Lee, V. E., & Loeb, S. (2000). School Size in Chicago Elementary Schools: Effects on Teachers' Attitudes and Students' Achievement. *American Educational Research Journal*, 37(1), 3-31.

Leithwood, K. A., & Earl, L. (2000). Educational Accountability Effects: An International Perspective. *Peabody Journal of Education*, 75(4), 1-18.

Leithwood, K. A., Louis, K. S., & Anderson, S. E. (2012). Linking leadership to student learning (pp. xxviii, 282 p.). Retrieved from <http://ezproxy.msu.edu:2047/login?url=http://site.ebrary.com/lib/michstate/Top?id=10514037>

Lewis, A. L. (2010). *School leaders as both colonized and colonizers: Understanding professional identity in an era of No Child Left Behind*. Unpublished doctoral dissertation, University of Illinois, Urbana Champaign, IL.

Libresco, A. S. (2005). How she stopped worrying and learned to love the test..., sort of. In E. A. Yeager & O. L. Davis (Eds.), *Wise social studies teaching in an age of high-stakes testing*

- (pp. 33-49). Greenwich, CT: Information Age Publishing.
- Lillard, D. R., & DeCicca, P. P. (2001). Higher Standards, More Dropouts? Evidence within and across Time. *Economics of Education Review*, 20(5), 459-473.
- Lind, V. R. (2007). High Quality Professional Development: An Investigation of the Supports for and Barriers to Professional Development in Arts Education. *International Journal of Education & the Arts*, 8(2), 1-18.
- Lindstrom, P. H., & Speck, M. (2004). *The principal as professional development leader*. Thousand Oaks, CA: Corwin Press.
- Linn, R. L. (2000). Assessments and Accountability. *Educational Researcher*, 29(2), 4-16.
- Lipsky, M. (2010). *Street-level bureaucracy : dilemmas of the individual in public services* (30th anniversary expanded ed.). New York, NY: Russell Sage Foundation.
- Loeb, S., Miller, L. C., & Strunk, K. O. (2009). The State Role in Teacher Professional Development and Education throughout Teachers' Careers. *Education Finance and Policy*, 4(2), 212-228.
- Louis, K. S., Febey, K., & Schroeder, R. (2005). State-mandated accountability in high schools: Teachers' interpretations of a new era. *Educational Evaluation and Policy Analysis*, 27(2), 177-204.
- Louis, K. S., Leithwood, K., Anderson, S. E., & Wahlstrom, K. L. (2010). *Learning from Leadership: Investigating the Links to Improved Student Learning. The Informed Educator Series*. Alexandria, VA: Educational Research Service.
- Louis, K. S., Marks, H. M., & Sharon, K. (1996). Teachers' Professional Community in Restructuring Schools. *American Educational Research Journal*, 33(4), 757-798.
- Louis, K. S., & Robinson, V. M. (2012). External mandates and instructional leadership: school leaders as mediating agents. *Journal of Educational Administration*, 50(5), 629-665.
- Luschei, T. F., & Christensen, G. S. (2008). District Leaders Eroding School Coherence? The Interpretation of Accountability Mandates. *Research in the Sociology of Education*, 16, 67-101.

- Maas, C. J. M., & Hox, J. J. (2002). Sample sizes for multilevel modeling.
- Maas, C. J. M., & Hox, J. J. (2004). The influence of violations of assumptions on multilevel parameter estimates and their standard errors. *Computational Statistics & Data Analysis*, 46(3), 427-440.
- Marks, H. M., & Nance, J. P. (2007). Contexts of Accountability Under Systemic Reform: Implications for Principal Influence on Instruction and Supervision. *Educational Administration Quarterly*, 43(1), 3-37.
- Marsh, D. D., & LeFever, K. (2004). School principals as standards-based educational leaders: Looking across policy contexts. *Educational Management Administration & Leadership*, 32(4), 387-404.
- Martell, C. (2010). *Continuously Uncertain Reform Effort: State-Mandated History and Social Science Curriculum and the Perceptions of Teachers*. Paper presented at the Paper presented at the Annual Meeting of the American Educational Research Association, Denver, Colorado.
- McDermott, K. A. (2003). What Causes Variation in States' Accountability Policies? *Peabody Journal of Education*, 78(4), 153-176.
- McDonald, D. (2002). Education Policy Shifts in New Law. *Momentum*, 33(2), 73-74.
- McDonnell, L. M. (2005). No Child Left Behind and the Federal Role in Education: Evolution or Revolution? *Peabody Journal of Education*, 80(2), 19-38.
- McGhee, M. W., & Nelson, S. W. (2005). Sacrificing Leaders, Villainizing Leadership: How Educational Accountability Policies Impair School Leadership. *Phi Delta Kappan*, 86(5), 367-372.
- McGuinn, P. (2012). Stimulating Reform: Race to the Top, Competitive Grants and the Obama Education Agenda. *Educational Policy*, 26(1), 136-159.
- McIntosh, S. (2012). *State High School Exit Exams: A Policy in Transition*. Center on Education Policy. 2140 Pennsylvania Avenue NW Room 103, Washington, DC 20037.
- McKay, R. W. (2011). *The Effect of No Child Left Behind on Elementary School Principals as*

*Instructional Leaders*. Unpublished doctoral dissertation, Walden University, Minneapolis, MN.

McLaughlin, D., Bandeira de Mello, V., Blankenship, C., Chaney, K., Esra, P., Hikawa, H., . . . Wolman, M. (2008). *Comparison Between NAEP and State Reading Assessment Results: 2003*. Washington, DC.: National Center for Education Statistics.

McNeil, L. M. (2000). *Contradictions of school reform : educational costs of standardized testing*. New York, NY: Routledge.

McNeil, M. (2011). Stimulus Aid Sparked Progress on Some Goals, States Say. *Education Week*, 30(32), 16-17.

Meyer, L., Orlofsky, G. F., Skinner, R. A., & Spicer, S. (2002). The state of the states. *Education Week*, 21(17), 68-70.

Mills, J. I. (2008). A Legislative Overview of No Child Left Behind. In T. Berry, R. M. & e. Eddy (Eds.), *Consequences of No Child Left Behind for educational evaluation* (pp. 9-20). San Francisco, CA: Jossey Bass. .

Mintrop, H. (2003). The Limits of Sanctions in Low-Performing Schools: A Study of Maryland and Kentucky Schools on Probation. *Education Policy Analysis Archives*, 11(3).

Mintrop, H., & Sunderman, G. L. (2009). Predictable Failure of Federal Sanctions-Driven Accountability for School Improvement--And Why We May Retain It Anyway. *Educational Researcher*, 38(5), 353-364.

Mojkowski, C. (2000). The essential role of principals in monitoring curriculum implementation. *National Association of Secondary School Principals. NASSP Bulletin*, 84(613), 76-83.

Murnane, R. J., & Levy, F. (2001). Will standards-based reforms improve the education of students of color? *National Tax Journal*, 54(2), 401-416.

Murnane, R. J., & Papay, J. P. (2010). Teachers' Views on No Child Left Behind: Support for the Principles, Concerns about the Practices. *The Journal of Economic Perspectives*, 24(3), 151-166.

Nance, J., & Marks, H. (2008). Curriculum and instruction policy in the context of multiple



- accountabilities. In W. K. Hoy & M. F. DiPaola (Eds.), *Improving schools : studies in leadership and culture* (pp. 193-221). Charlotte, NC: Information Age Pub.
- National Center for Education Statistics, E. (2007). *Mapping 2005 State Proficiency Standards onto the NAEP Scales*. Washington, DC: National Center for Education Statistics.
- National Center for Education Statistics, E. (2011). *Mapping State Proficiency Standards onto the NAEP Scales: Variation and Change in State Standards for Reading and Mathematics, 2005-2009. NCES 2011-458*. Washington, DC: National Center for Education Statistics. .
- Neal, D., & Schanzenbach, D. W. (2010). Left behind by design: Proficiency counts and test-based accountability. *The Review of Economics and Statistics*, 92(2), 263-283.
- Nelson, S. W., De la Colina, M. G., & Boone, M. D. (2008). Lifeworld or systemworld: what guides novice principals? *Journal of Educational Administration*, 46(6), 690-701.
- Newmann, F. M., King, M. B., & Youngs, P. (2000). Professional development that addresses school capacity: Lessons from urban elementary schools. *American Journal of Education*, 108(4), 259-299.
- Nichols, S. L., & Berliner, D. C. (2007). *Collateral damage : how high-stakes testing corrupts America's schools*. Cambridge, MA: Harvard Education Press.
- Nichols, S. L., Glass, G. V., & Berliner, D. C. (2006). High-Stakes Testing and Student Achievement: Does Accountability Pressure Increase Student Learning? *Education Policy Analysis Archives*, 14(1), 1-172.
- O'Day, J. A. (2002). Complexity, accountability, and school improvement. *Harvard Educational Review*, 72(3), 293-329.
- O'Donnell, R., & White, G. (2005). Within the Accountability Era: Principals' Instructional Leadership Behaviors and Student Achievement. *National Association of Secondary School Principals. NASSP Bulletin*, 89(645), 56-71.
- O'Hara, D. P. (2006). Teacher autonomy: Why do teachers want it, and how do principals determine who deserves it? [3227718]. *Ed.D.*, 143 p.
- Oleszewski, A., Shoho, A., & Barnett, B. (2012). The development of assistant principals: a

- literature review. *Journal of Educational Administration*, 50(3), 264-286.
- Opdycke, W. S. (2004). *An analysis of the elementary principal's role in implementing school accountability within California's High Priority School: A case study*. Unpublished doctoral dissertation, University of Southern California, Los Angeles, CA.
- Owings, W. A., Kaplan, L. S., & Chappell, S. (2011). Troops to Teachers as School Administrators: A National Study of Principal Quality. *National Association of Secondary School Principals. NASSP Bulletin*, 95(3), 212-236.
- Pearson, L. C. (1995). The prediction of teacher autonomy from a set of work-related and attitudinal variables. *Journal of Research & Development in Education*, 28(2), 79-85.
- Pearson, L. C., & Hall, B. W. (1993). Initial Construct Validation of the Teaching Autonomy Scale. *Journal of Educational Research*, 86(3), 172-178.
- Pearson, L. C., & Moomaw, W. (2005). The Relationship between Teacher Autonomy and Stress, Work Satisfaction, Empowerment, and Professionalism. *Educational Research Quarterly*, 29(1), 37-53.
- Pearson, L. C., & Moomaw, W. (2006). Continuing Validation of the Teaching Autonomy Scale. *The Journal of Educational Research*, 100(1), 44-51,64.
- Pedulla, J. J., Abrams, L. M., Madaus, G. F., Russell, M. K., Ramos, M. A., & Miao, J. (2003). *Perceived Effects of State-Mandated Testing Programs on Teaching and Learning: Findings from a National Survey of Teachers*. Chestnut Hill, MA: Lynch School of Education, Boston College.
- Peterson, P. E., & Hess, F. M. (2005). Johnny Can Read...in Some States: Assessing the Rigor of State Assessment Systems. *Education Next*, 5(3), 52-53.
- Phillips, K. J., Desimone, L., & Smith, T. M. (2011). Teacher Participation in Content-Focused Professional Development & the Role of State Policy. *Teachers College Record*, 113(11), 2586-2621.
- Polk, S. (2006). *Creating school cultures that support strong internal accountability systems: "It's all about the relationships". A study of leadership at two Chicago charter schools*. Ed.D., Harvard University, United States -- Massachusetts.

- Porter, A. C., Linn, R. 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.
- Printy, S. M. (2010). How principals influence instructional practice. In W. K. Hoy & M. F. DiPaola (Eds.), *Analyzing school contexts : influences of principals and teachers in the service of students* (pp. 71-102). Charlotte, NC: Information Age Pub.
- Priolo, G. R. (2010). *How principals lead in an era of testing and accountability: A qualitative study of the perceptions of principals leading schools on the continuum of No Child Left Behind sanctions*. Unpublished doctoral dissertation, Temple University, Philadelphia, PA.
- Putnam, R. T., & Borko, H. (1997). Teacher learning: Implications of new views of cognition. In B. J. Biddle, T. L. Good & I. Goodson (Eds.), *International handbook of teachers and teaching*. Dordrecht ; Boston: Kluwer Academic Publishers.
- Quiocho, A., & Stall, P. (2008). NCLB and Teacher Satisfaction. [Article]. *Leadership*, 37(5), 20-24.
- Range, B. G., Scherz, S., Holt, C. R., & Young, S. (2011). Supervision and evaluation: The Wyoming perspective. *Educational Assessment, Evaluation and Accountability*, 23(3), 243-265.
- Raudenbush, S., W. , & Bryk, A. S. (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, CA: Sage Publications, Inc.
- Ravitch, D. (2002). A brief history of testing and accountability. *Hoover Digest*, 4.
- Ravitch, D. (2010). We've Always Had National Standards, *Education Week* Retrieved from <http://www.edweek.org/ew/articles/2010/01/14/17ravitch-comm.h29.html>
- Reback, R. (2008). Teaching to the rating: school accountability and the distribution of student achievement. *Journal of Public Economics*, 92(5-6), 1394-1415.
- Reback, R., Rockoff, J., & Schwartz, H. L. (2011). *Under Pressure: Job Security, Resource Allocation, and Productivity in Schools under NCLB*. Cambridge, MA: National Bureau of Economic Research.

- Reed, E., Scull, J., Slicker, G., & Winkler, A. M. (2012). *Defining Strong State Accountability Systems: How Can Better Standards Gain Greater Traction? A First Look*. Washington, DC: Thomas B. Fordham Institute.
- Rentner, D. S., Scott, C., Kober, N., Chudowsky, N., Chudowsky, V., Joftus, S., & Zabala, D. (2006). *From the capital to the classroom: Year 4 of the No Child Left Behind Act*. Washington, DC: Center on Education Policy.
- Resnick, D. P. (1980). Minimum Competency Testing Historically Considered. *Review of Research in Education*, 8(1), 3-29.
- Rex, L. A., & Nelson, M. C. (2004). How Teachers' Professional Identities Position High-Stakes Test Preparation in Their Classrooms. *Teachers College Record*, 106(6), 1288-1331.
- Rice, J. K. (2010). *Principal Effectiveness and Leadership in an Era of Accountability: What Research Says*. Washington, DC: National Center for Analysis of Longitudinal Data in Education Research, The Urban Institute.
- Rice, J. K., & Malen, B. (2003). The Human Costs of Education Reform: The Case of School Reconstitution. *Educational Administration Quarterly*, 39(5), 635-666.
- Robinson, V. M. J., Lloyd, C. A., & Rowe, K. J. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly*, 44(5), 635-674.
- Roellke, C., & Rice, J. K. (2008). Responding to Teacher Quality and Accountability Mandates: The Perspective of School Administrators and Classroom Teachers. *Leadership and Policy in Schools*, 7(3), 264-295.
- Rorrer, A. K., & Skrla, L. (2005). Leaders as Policy Mediators: The Reconceptualization of Accountability. *Theory Into Practice*, 44(1), 53-62.
- Roth, G., Assor, A., Kanat-Maymon, Y., & Kaplan, H. (2007). Autonomous motivation for teaching: How self-determined teaching may lead to self-determined learning. *Journal of Educational Psychology*, 99(4), 761-774.
- Rothman, R. (1995). *Measuring up : standards, assessment, and school reform* (1st ed.). San Francisco, CA: Jossey-Bass.

- Rothstein, R. (2004). *Class and schools : using social, economic, and educational reform to close the black-white achievement gap*. New York, NY: Economic Policy Institute: Teachers College, Columbia University.
- Rothstein, R., Jacobsen, R., & Wilder, T. (2008). The Outcome Goals of American Public Education. In R. Rothstein, R. Jacobsen & T. Wilder (Eds.), *Grading Education: Getting Accountability Right* (pp. 13-34). Washington, DC: Economic Policy Institute.
- Rouse, C. E., Hannaway, J., Goldhaber, D., & Figlio, D. (2007). *Feeling the Florida Heat? How Low-Performing Schools Respond to Voucher and Accountability Pressure*. Washington, DC: The Urban Institute.
- Rudolph, L. (2006). *Decomposing teacher autonomy: A study investigating types of teacher autonomy and how it relates to job satisfaction*. Unpublished doctoral dissertation University of Pennsylvania, Philadelphia, PA.
- Rutledge, S. A., Harris, D. N., & Ingle, W. K. (2010). How Principals "Bridge and Buffer" the New Demands of Teacher Quality and Accountability: A Mixed-Methods Analysis of Teacher Hiring. *American Journal of Education*, 116(2), 211-242.
- Sanzo, K. L., Sherman, W. H., & Clayton, J. (2011). Leadership Practices of Successful Middle School Principals. *Journal of Educational Administration*, 49(1), 31-45.
- Sato, M., Wei, R. C., & Darling-Hammond, L. (2008). Improving Teachers' Assessment Practices Through Professional Development: The Case of National Board Certification. *American Educational Research Journal*, 45(3), 669-700.
- Schmidt, M. I., & Datnow, A. (2005). Teachers' sense-making about comprehensive school reform: The influence of emotions. *Teaching and Teacher Education*, 21(8), 949-965.
- Schneider, M. (2011). *The Accountability Plateau*. Washington, DC: Thomas B. Fordham Institute.
- Scott, G. A. (2011). *Race to the Top: Reform Efforts Are Under Way and Information Sharing Could Be Improved. Report to Congressional Committees. GAO-11-658*. Washington, DC US Government Accountability Office.
- Scott, W. R., & Meyer, J. W. (1983). The Organization of Societal Sectors. In J. W. Meyer & W. R. Scott (Eds.), *Organizational environments : ritual and rationality*. Beverly Hills: Sage.

- Sebastian, J., & Allensworth, E. (2012). The Influence of Principal Leadership on Classroom Instruction and Student Learning: A Study of Mediated Pathways to Learning. *Educational Administration Quarterly*, 48(4), 626-663.
- Sebring, P. B., & Bryk, A. S. (2000). School Leadership and the Bottom Line in Chicago. *Phi Delta Kappan*, 81(6), 440-443.
- Shelly, B. (2012). Flexible Response: Executive Federalism and the No Child Left Behind Act of 2001. *Educational Policy*, 26(1), 117-135.
- Shen, J. (1997). Has the Alternative Certification Policy Materialized Its Promise? A Comparison between Traditionally and Alternatively Certified Teachers in Public Schools. *Educational Evaluation and Policy Analysis*, 19(3), 276-283.
- Shen, J., & Xia, J. (2012). The Relationship between Teachers' and Principals' Decision-Making Power: Is It a Win-Win Situation or a Zero-Sum Game? *International Journal of Leadership in Education*, 15(2), 153-174.
- Shepard, L. (2008). A Brief history of accountability testing, 1965-2007. In K. E. Ryan & L. A. Shepard (Eds.), *The future of test-based educational accountability* (pp. 25-46). New York: NY: Routledge.
- Shih, T.-H. (2008). *Adequate sample sizes for viable 2-level hierarchical linear modeling analysis: A study on sample size requirement in HLM in relation to different intraclass correlations*. 3302215, University of Virginia, United States -- Virginia.
- Shuster, K. (2012). Re-Examining Exit Exams: New Findings from the Education Longitudinal Study of 2002. *Education Policy Analysis Archives*, 20(3), 35-35.
- Sirotnik, K. A., & Kimball, K. (1999). Standards for Standards-Based Accountability Systems. *Phi Delta Kappan*, 81(3), 209-214.
- Smith, J. B., Smith, B., & Bryk, A. S. (1998). *Setting the Pace: Opportunities To Learn in Chicago's Elementary Schools. Improving Chicago's Schools*. Consortium on Chicago School Research, 1313 E. 60th Street, Chicago, IL 60637 (\$10).
- Smith, S. S., & Mickelson, R. A. (2000). All that glitters is not gold: School reform in Charlotte-Mecklenburg. *Educational Evaluation and Policy Analysis*, 22(2), 101-127.

- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: an introduction to basic and advanced multilevel modeling* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Sparks, D. (2012). *The relationship between teacher perceptions of autonomy in the classroom and standards based accountability reform*. Ph.D., University of Maryland, College Park, Ann Arbor.
- Spicer, K. A. (2008). *Professional development in the era of accountability: Teacher's perceptions*. Unpublished doctoral dissertation, University of Virginia, Charlottesville, VA.
- Spillane, J. P., Diamond, J. B., Burch, P., & Hallett, T. (2002). Managing in the middle: School leaders and the enactment of accountability policy. *Educational Policy*, 16(5), 731-762.
- Spillane, J. P., Parise, L. M., & Sherer, J. Z. (2011). Organizational routines as coupling mechanisms: Policy, school administration, and the technical core. *American Educational Research Journal*, 48(3), 586-619.
- Spiri, M. H. (2001). *School Leadership and Reform: Case Studies of Philadelphia Principals. Occasional Paper*. Philadelphia, PA Consortium for Policy Research in Education
- Springer, M. G. (2008). The Influence of an NCLB Accountability Plan on the Distribution of Student Test Score Gains. *Economics of Education Review*, 27(5), 556-563.
- Srikantaiah, D. (2009). *How State and Federal Accountability Policies Have Influenced Curriculum and Instruction in Three States: Common Findings from Rhode Island, Illinois, and Washington*. Washington, DC: Center on Education Policy.
- Stecher, B. M., Chun, T., Barron, S., & Ross, K. (2000). *The Effects of the Washington State Education Reform on Schools and Classrooms: Initial Findings*. Santa Monica, CA: The Rand Corporation.
- Steffy, B. E. (2000). *Life cycle of the career teacher*. Indianapolis, IN: Thousand Oaks, Calif.
- Stevenson, D. L., & Schiller, K. S. (1999). State Education Policies and Changing School Practices: Evidence from the National Longitudinal Study of Schools, 1980-1993. *American Journal of Education*, 107(4), 261-288.

- Stevenson, K. R. (2006). *School Size and Its Relationship to Student Outcomes and School Climate: A Review and Analysis of Eight South Carolina State-Wide Studies*. Washington, DC National Clearinghouse for Educational Facilities.
- Sunderman, G. L. (2006). *The Unraveling of No Child Left Behind: How Negotiated Changes Transform the Law*. Cambridge, MA: Harvard Education Publishing Group.
- Sunderman, G. L., Orfield, G., & Kim, J. S. (2006). FLAWED ASSUMPTIONS: How No Child Left Behind Fails Principals. *Principal Leadership*, 6(8), 16-19.
- Swanson, C. B., & Stevenson, D. L. (2002). Standards-Based Reform in Practice: Evidence on State Policy and Classroom Instruction from the NAEP State Assessments. *Educational Evaluation and Policy Analysis*, 24(1), 1-27.
- Tannen, D. (1991). *You just don't understand: women and men in conversation* (1st Ballantine Books PBK. ed.). New York, NY: Ballantine.
- Taylor, J., Stecher, B., O'Day, J., Naftel, S., & Le Floch, K. C. (2010). *State and Local Implementation of the "No Child Left Behind Act". Volume IX--Accountability under "NCLB": Final Report*. Jessup, MD: US Department of Education.
- Theall, K. P., Scribner, R., Broyles, S., Yu, Q., Chotalia, J., Simonsen, N., . . . Carlin, B. P. (2011). Impact of small group size on neighbourhood influences in multilevel models. *Journal of Epidemiology and Community Health*, 65(8), 688-695.
- Togneri, W., & Anderson, S. E. (2003). *Beyond Islands of Excellence: What Districts Can Do To Improve Instruction and Achievement in All Schools. A Project of the Learning First Alliance [and] A Leadership Brief*. Washington, DC: Learning First Alliance.
- Usher, A. (2012). *AYP Results for 2010-11*. Washington, DC: Center on Education Policy.
- Vogler, K. E. (2008). Comparing the impact of state accountability examinations on Mississippi and Tennessee social studies teachers' instructional practices. *Educational Assessment*, 13(1), 1-32.
- Wagner, R. B. (1989). *Accountability in education : a philosophical inquiry*. New York, NY: Routledge.



- Wahlstrom, K. L., & York-Barr, J. (2011). Leadership: Support and Structures Make the Difference for Educators and Students. *Journal of Staff Development*, 32(4), 22-25.
- Walberg, H. J. (2002). *School Accountability: Principles for Accountability Designs*. Stanford, CA: Hoover Institution Press.
- Wei, X. (2008). *Accountability stringency, incentives and student performance*. Unpublished doctoral dissertation, Stanford University, Stanford, CA.
- Wei, X. (2012). Are More Stringent NCLB State Accountability Systems Associated With Better Student Outcomes? An Analysis of NAEP Results Across States. *Educational Policy*, 26(2), 268-308.
- West, M. (2007). *Testing, Learning, and Teaching: The Effects of Test-Based Accountability on Student Achievement and Instructional Time in Core Academic Subjects*. Washington, DC: The Thomas B. Fordham Institute.
- Whitener, E. M. (2001). Do 'high commitment' human resource practices affect employee commitment? A cross-level analysis using hierarchical linear modeling. *Journal of Management*, 27(5), 515-535.
- Wills, J. S., & Sandholtz, J. H. (2009). Constrained Professionalism: Dilemmas of Teaching in the Face of Test-Based Accountability. *Teachers College Record*, 111(4), 1065-1114.
- Winters, M. A., Trivitt, J. R., & Greene, J. P. (2010). The Impact of High-Stakes Testing on Student Proficiency in Low-Stakes Subjects: Evidence from Florida's Elementary Science Exam. *Economics of Education Review*, 29(1), 138-146.
- Wirt, F. M., & Kirst, M. W. (1989). *Schools in conflict : the politics of education* (2nd ed.). Berkeley, Calif.: McCutchan Pub. Corp.
- Wise, A. E. (1979). *Legislated learning : the bureaucratization of the American classroom*. Berkeley: University of California Press.
- Wollman-Bonilla, J. E. (2004). Principled Teaching to(wards) the Test?: Persuasive Writing in Two Classrooms. *Language Arts*, 81(6), 502-511.
- Yeh, S. S. (2005). Limiting the Unintended Consequences of High-Stakes Testing. *Education*

*Policy Analysis Archives*, 13(43).

Yoon, K. S., Duncan, T., Lee, S., Shapley, K., Scarloss, B., Taylor, J., . . . Streke, A. (2008). *The effects of teachers' professional development on student achievement: Findings from a systematic review of evidence*. Paper presented at the Paper presented at the Annual Meeting of the American Educational Research Association, New York, NY.

Youngs, P., & King, M. B. (2002). Principal leadership for professional development to build school capacity. *Educational Administration Quarterly*, 38(5), 643-670.

Zabala, D., Minnici, A., McMurrer, J., & Briggs, L. (2008). *State High School Exit Exams: A Move toward End-of-Course Exams*. Washington, DC: Center on Education Policy.

Zabala, D., Minnici, A., McMurrer, J., Hill, D., Jennings, J., & Kober, N. (2007). *State High School Exit Exams: Working to Raise Test Scores*. Washington, DC: Center on Education Policy.