

ABSTRACT

Title of Document: THE RELATOINSHIP BETWEEN TEACHER PERCEPTIONS OF AUTONOMY IN THE CLASSROOM AND STANDARDS BASED ACCOUNTABILITY REFORM

Dina Sparks, Doctor of Philosophy, 2012

Directed By: Dr. Robert Croninger and Dr. Betty Malen,
Department of Teaching and Learning, Policy
and Leadership

Over the past 30 years, standards based accountability reform (SBA) has taken hold in public education. SBA reform includes defined academic expectations, curricula standards, measureable assessments, and performance accountability. SBA impacts multiple levels of public education. Its most recent federal codification, the No Child Left Behind (NCLB) Act, includes sanctions meant to influence what happens in classrooms. Historically, teachers have held a great deal of control over the activities in the classroom. Research suggests that teacher control (i.e. autonomy) over the classroom often resulted in uneven implementation of reform policies across schools, the transformation of policies to fit existing practice or the insulation of classrooms altogether from policy reform. To achieve its stated goals, SBA seeks to influence teacher and school practices, particularly where students fail to meet performance goals.

This study examines the intersection of teacher perceptions of autonomy and SBA reforms, including NCLB. The study uses four waves of nationally representative Schools and Staffing Survey data from 1993-94 to 2007-08 to investigate changes in teacher autonomy over time and to examine specific school and teachers characteristics associated with changes in autonomy in 2007-08.

Over-time findings reveal that teachers perceived lower classroom autonomy between 2003-04 and 2007-08. Across all four waves of data, the variation in teachers' classroom autonomy increased, and more of this increased variation occurred between schools rather than within schools. Findings for 2007-08 reveal that teachers who taught

in elementary schools or taught tested subjects perceived lower levels of autonomy than did teachers in secondary schools or who taught non-tested subjects. Further analyses based on state application of adequate yearly progress (AYP) sanctions revealed a differential effect on teacher autonomy for Title I schools and for schools that failed to make AYP. Findings from this study suggest that although NCLB targets Title I schools, teachers in all schools perceive lower autonomy based on the grade level and the subject matter taught, and that state policies regarding NCLB may lead to uneven or unintended effects on teacher perceptions of autonomy in the classroom.

THE RELATIONSHIP BETWEEN TEACHER PERCEPTIONS OF
AUTONOMY IN THE CLASSROOM AND
STANDARDS BASED ACCOUNTABILITY REFORM

by

Dina Sparks

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2012

Advisory Committee:

Dr. Robert Croninger, Associate Professor, Chair
Dr. Betty Malen, Professor, Co-Chair
Dr. Jennifer Rice, Professor
Dr. Linda Valli, Professor
Dr. Margaret McLaughlin, Professor

© Copyright by
Dina Sparks
2012

Table of Contents

Table of Contents	ii
List of Figures.....	v
List of Tables	vi
Chapter 1: Introduction	1
Defining teacher autonomy.....	2
SBA theory of action	4
Relationship Between Teacher Autonomy and SBA.....	6
Conceptual Framework.....	10
Research Questions	16
Overall Design of the Study.....	18
Potential Significance	18
Chapter 2: Literature Review.....	21
Historical Context of SBA Reform Development and Implementation	21
Elements of SBA Reform Policies.....	25
High stakes and low stakes accountability.....	26
The growth of the federal role in education.....	27
SBA Implementation across the public school system	34
Federal level.....	34
State level.....	40
District level.....	43
School level.....	45
Research on Teacher Autonomy	48
Teacher autonomy and professionalism.....	49
Teacher autonomy and job satisfaction.....	51
Teacher autonomy and retention.....	54
Teacher autonomy and SBA reform.	61
Implications of Literature on Teacher Autonomy for the Present Study.....	71

Chapter 3: Methods	73
Description of Data Source	73
School file.	76
Principal file.	76
Teacher file.	76
Empirical Frameworks	76
Analytic Samples	82
Outcome variable.	85
Scale	87
Factor	89
Indicator and control variables.	91
Weighting variables	92
Statistical Procedures	96
Analyses of variation.	96
ICC comparisons.	98
Hierarchical Linear Model (HLM) analysis.	99
Limitations of the Study.	101
Chapter 4: Results.....	104
Over-Time Analyses: 1993–94, 1999–2000, 2003–04 and 2007–08	104
Full Sample Analyses: 2007–08	112
Descriptive statistics.	113
Multilevel model results.	115
Teacher and school characteristics.	118
Model 1.	118
Model 2.	119
Model 3.	123
AYP status.	125
Subsample analyses: 2007–08	126
Descriptive results.....	126
Multilevel model results	128

Summary of results	132
Chapter 5: Discussion	134
Discussion of the over-time hypothesis	134
Discussion of the differential hypotheses	136
Tested subjects and teacher-perceived autonomy.....	137
Experience levels and teacher-perceived autonomy.	139
School leadership and teacher-perceived autonomy.....	140
Title I status and teacher-perceived autonomy.	142
AYP Status and teacher-perceived autonomy.....	143
Concluding remarks.	147
Implications for Future Research.....	149
Appendix.....	152
References	161

List of Figures

Figure 1. Relationship between standards-based reform and teacher perceptions of scope and degree of autonomy in the classroom domain.....	12
Figure 2. Empirical model of change in degree of teacher-perceived autonomy in the classroom domain: 1993—2007.	79
Figure 3. Empirical model of the relationships between SBA reform, school and teacher characteristics, and the degree of teacher-perceived autonomy in the classroom domain.....	80
Figure 4. Autonomy scale mean values across teachers: 1993–94, 1999–00, 2003–04, 2007–08.....	105
Figure 5. Coefficients of variation: 1993–94, 1999–2000, 2003–04, 2007–08.....	107
Figure 6. ICC over time: 1999–93, 1999–00, 2003–04, 2007–08.	109
Figure 7. ICC over time by Title I status: 1993–94, 1999–00, 2003–04, 2007–08.	110

List of Tables

Table 1	AYP sanctions.....	36
Table 2	Sample size	75
Table 3	Unweighted analytic sample sizes by year	83
Table 4	School-level missing data analysis	84
Table 5	Variables included in teacher autonomy construct.....	86
Table 6	Cronbach alphas.....	89
Table 7	Description of variables used.....	94
Table 8	Descriptive statistics for the analytic sample: School year 2007-08	114
Table 9	Title I status by AYP status: School year 2007-08.....	115
Table 10	Multilevel models for scope and degree of teacher-perceived autonomy in the classroom domain ^a	117
Table 11	Experience effect on novice status for model 2.....	122
Table 12	Experience and interaction effects on novice teachers in elementary and secondary schools for model 3.....	124
Table 13	Descriptive statistics for subsamples by state sanction policies: School year 2007-08	127
Table 14	Multilevel models for scope and degree of teacher-perceived autonomy in the classroom domain by state-level sanction policy ^a	130

Chapter 1: Introduction

Decades of research on education reform implementation suggest that as street-level bureaucrats (Lipsky, 1980) educators in general and teachers in particular have had the ability to insulate themselves and mediate external intervention in the classroom, either by keeping reform from permeating the classroom or by adapting policies to fit established practice (Allensworth, Ponisciak, & Mazzeo, 2009; Cohen, Moffitt, & Goldin, 2007; Elmore, 1987; Ingersoll & May, 2010, 2011; Rowan, 1990; Spillane & Zeuli, 1999; Weatherly & Lipsky, 1977). Cuban (1990) discusses instruction, curriculum, and centralization policies as examples of recurring school reforms that illustrate the challenge of federal, state, and district policies “ever get[ting] past the classroom door” (p. 3). He points out that implementation of these policies ultimately rests with teachers who are decoupled from administrators and policymakers and are able to “ignore and modify” (p. 11) policies. Meanwhile, principals and teachers tend to operate under informal agreements whereby both parties maintain separate spheres of influence within schools with little to no interference from one another (Hanson, 2006).

Currently, questions remain as to whether educators continue to have the ability to insulate their workplace given the standards-based accountability reforms (SBA) that have enveloped public education in the United States. The SBA reforms define academic expectations, curricula standards, measureable assessments, and performance accountability systems (Hamilton, Stecher, & Yuan, 2008). Ostensibly, SBA reforms outline a systemic structure for consistent and quality public education. In reality, SBA reforms challenge the status quo of education delivery and the freedom that permitted

educators to ignore or modify externally mandated interventions at the school and classroom levels. Specifically, SBA reforms break the historic *détente* between teachers and other school authorities and force a rebalancing of external demands on the education system and internal practices of classroom teachers.

Defining teacher autonomy

One way of understanding SBA reforms is as an attempt to restrict teacher autonomy, particularly the autonomy of teachers who fail to meet performance goals or comply with external mandates. What constitutes teacher autonomy, however, can be difficult to define. Scholars have used a variety of terms, such as *curriculum control* (Archbald & Porter, 1994) or *professional capacity* (Cohen, 1996), to describe teacher autonomy. Nonetheless, for scholars “the concept of professional autonomy is unruly and mired with multiple, ambiguous interpretations” (Boote, 2006, p. 462). Ingersoll (2003) defines teacher autonomy as a dimension of power that is “a function of the extent to which [teachers] influence the decisions that are most central to their work” (p. 47). He uses the terms *power* and *control* interchangeably when he investigates teacher autonomy in schools and declares teachers to be “short on power, long on responsibility” when describing their influence in schools (Ingersoll, 2007, p. 20). While the literature offers no absolute definition of teacher autonomy, the current study incorporates aspects of prominent definitions that draw on concepts of freedom, discretion, and power over academic and social aspects of schooling to define teacher autonomy in the classroom (Bogler & Somech, 2004; Firestone & Pennell, 1993; Ingersoll, 2003). Specifically, I define teacher autonomy as the power to exercise control over key elements of work in the classroom.

I do not adopt the notion that the power to exercise control over key elements of work in the classroom is synonymous with power over all decisions that enter the classroom, or extend the definition of teacher autonomy beyond the classroom. Rather, I acknowledge that teacher autonomy is “situationally constrained” (Cuban, 1990, p. 11) through structural and administrative elements of schools that condition the exercise of control in specific domains. Similar to Ingersoll, authors have used a host of terms as indicative of or equivalent to the degree of teacher autonomy in the classroom. I do not attempt to disentangle this “semantic thicket” (Geary, 1992) but offer a functional definition of the teacher autonomy construct. In this dissertation, the study of autonomy is the study of the scope and degree of control that teachers have over key elements of their work in the classroom domain.

Autonomy in general and teacher autonomy in particular are important constructs. Social psychologists document the general positive relationship between autonomy and job satisfaction (Deci & Ryan, 1987; Richer, Blanchard, & Vallerand, 2002). Parallel to social psychologists, education researchers document that teacher autonomy associates generally with several aspects of workplace conditions, including teacher job satisfaction, teacher motivation, and teacher retention (Guarino, Santibanez, & Daley, 2006; Ingersoll, 2003; Ingersoll & May, 2011; Kim & Loadman, 1994; Pearson, 1995; Pearson & Moomaw, 2006). Further, teacher autonomy is a policy-amenable construct that can affect teachers’ working conditions depending on how policymakers and administrators develop, frame, and implement policy. For example, teachers may perceive policies that emphasize high-stakes accountability in schools as “an imposition on their professional autonomy” (Boyd, Grossman, Ing, Lankford, Loeb, & Wyckoff, 2011, p. 90) or they may

perceive them as policies that enable teachers to focus on core academic areas (Guggino & Brint, 2010). That educators may have opposing interpretations of SBA reform reveals a tension between teacher autonomy and school accountability that is inherent in the theory of action that underlies SBA policies, especially later forms as exemplified by No Child Left Behind (NCLB).

SBA theory of action

In the most general terms, SBA seeks to set defined and measureable student performance goals, assess student performance in light of those goals, and engender classroom practices that achieve those stated goals (Elmore, 2002; Hamilton, 2003; Stecher et al. 2008). To that end, SBA policies rely heavily on various combinations of incentives and sanctions to motivate educators to comply with policy provisions and adopt practices that advance the stated aims of SBA reforms. SBA assumes that these inducements and punishments are essential and effective levers to alter the behavior of educators in public school systems. While the stated goals of SBA reforms are to strengthen standards and accountability for all schools, intervention mechanisms (e.g., district supports, professional development) focus on schools that fail to meet academic performance targets and on teachers who have not successfully improved student academic achievement.

Beginning in the late 1980s, state- and district-level policymakers introduced SBA reform and states began to adopt SBA policies. Research on the early state-adopted SBA policies indicated that these policies “lead to behavioral change among teachers, [that] included increased scrutiny of test scores, targeted adjustments in teaching practice, closer alignment of curriculum with standards and assessments...and collaboration with

fellow teachers to improve performance” (Hannaway & Woodroffe, 2003). Elmore (2002) describes how “in early stages of the accountability movement, reformers had an expansive view of performance that included, in addition to tests, portfolios of student’s work, teachers’ evaluations of their students, student-initiated projects, and formal exhibitions of students’ work” (p. 32). However, in the late 1990s, reforms began to shift toward a test-centered approach to SBA reform as policymakers linked SBA policies to standardized test scores (Hamilton, 2003) Over time, SBA became synonymous with test-based accountability.

Federal policymakers embraced the appeal of test-based accountability when they codified testing through the No Child Left Behind Act (NCLB) in the 2001 reauthorization of the Elementary and Secondary Education Act (ESEA). Their underlying theory of action posited that the focus on education standards measured by test-based assessments and accompanied by sanctions for sub-par performance would prompt educators to adopt practices that would improve student performance (as measured by standardized test scores). Presumably the Title I school specific sanctions triggered by low student performance would motivate teachers, principals, and parents to make changes both within and beyond the classroom to enhance student performance and achieve “adequate yearly progress” (Hamilton et al., 2007; Hammond, 2007; Sunderman, Tracey, Kim & Orfield, 2004). Thus, NCLB includes: a mandatory benchmark for student performance in the form of Adequate Yearly Progress (AYP); a universal provision for public dissemination of a school’s AYP status (i.e., met or failed); a system-wide requirement for annual assessments in grades 3 through 8 and in mathematics, science, and English/Language Arts; targeted opportunities of school choice and

supplemental education services for students who attend schools that receive federal funding through ESEA's Title I program (i.e., Title I schools) and fail to make AYP two consecutive years; and specific actions toward school reconstitution for Title I schools that fail to make AYP more than two consecutive years.

Given the theory of action in SBA, particularly as adopted by NCLB, one might expect teachers of the NCLB-targeted grades, subjects and schools to change their classroom practices. For example, teachers might alter the content they emphasize, the teaching techniques they employ, and the approaches they use to manage the classroom and to prepare students for the targeted testing by subject and grade. One might also expect that teachers might perceive the pressure to alter aspects of their work in classrooms in at least two ways: as an infringement on their autonomy in the classroom or as an opportunity to offer additional support to historically lower-performing student populations in Title I schools (Sunderman, Tracey, Kim & Orfield, 2004). It is this intersection between SBA policies and teacher autonomy in the classroom that is central to the current study.

Relationship Between Teacher Autonomy and SBA

Generations of sociologists, economists, organizational theorists, and political scientists have used an array of theoretical approaches to study broad concepts of power, autonomy and control in organizations (Geary, 1992; Malen & Ogawa, 1988, Winter) and to study tensions between autonomy and accountability in the school setting (Booher-Jennings, 2005; Campbell, R. 1976; Fuhrman & Elmore, 1990; Hanushek & Raymond, 2002; Ingersoll, 2003, 2007; Manna, 2006a; Mintrop & Sunderman, 2009; Rowan, 1990). For example, researchers have framed studies on teaching as (a) a matter of technology,

where teaching is viewed as a “routinized technology” that can be managed through “control-based” organizational design; (b) or as a complex and non-routine technology, where teachers seek “authority, variety, and collegiality” in schools (Rowan, 1990); (c) as a matter of capacity, where through teacher education programs and professional development teachers are able to help students acquire knowledge and skills that lead to specified outcomes (Minnici, Hill, & Kober, 2007; Stecher et al., 2008); and (d) as a matter of profession where teachers are curricular and instructional experts charged with developing student intellects and citizenship (E. Campbell, 2006; Darling-Hammond, Berry, & Thoreson, 2001; Ingersoll, 2003).

Using various methodological approaches (e.g., case study, survey, statistical modeling) scholars have found that teachers’ perceptions of their autonomy affect their motivation to invest in teaching, their sense of job satisfaction, their view of teaching as a profession, and ultimately, their retention (Ingersoll & Alsalam, 1997; Ingersoll & May, 2011; Kim & Loadman, 1994; Pearson & Moomaw, 2005). Some scholars argue that as SBA implementation continues, “the distribution of power, authority, and control in schools is one of the most important issues in contemporary education research and policy” (Ingersoll & Merrill, 2011, p. 191). For example, in 2011 a study of New York City teachers sought to examine the relationship between contextual school characteristics (e.g., administrative support, school facilities) and teacher influence over their work and their retention. The authors found that teachers who perceived themselves to have low levels of autonomy had less influence over their work were more likely to either transfer to another school or leave the profession altogether (Boyd et al., 2011).

The amount of control over the classroom that teachers share with federal, state, and local actors affects the nature and degree of teacher autonomy. Historically, federal and state policymakers have viewed teacher autonomy as necessary for, and yet at times contrary to, the efficient and effective operation of schools. In the early 1990s, some reformers attempted to *increase* autonomy at the school level through site-based management and school improvement programs in which teachers shared control over the classroom with other local actors (i.e., parents) (White, 1992). However, within a decade, reformers turned toward more centralized controls that in effect sought to *decrease* autonomy at the school level (Glazer, 2008). SBA reforms represent the turn towards centralized strategies that change the distribution of power in schools, which may influence, for good or for ill, teacher autonomy.

Ingersoll (2003) states that “the distribution of control and influence in schools profoundly affects how well schools function” (p. 13). Thus, it is not surprising that the amount of power teachers have to control their work in the classroom is contested terrain. As SBA reforms take hold, policymakers have differing views of the effects SBA reforms may have on teacher autonomy depending on whether policymakers value high or low levels of control over teacher actions in the classroom. The polemic concerns at least two normative questions—Should teachers have power over the classroom? And if so, how much?—and two empirical questions—How much power do teachers have over the classroom? And what degree of teacher power over the classroom is associated with desirable education outcomes?¹

¹ The current study does not address the normative questions; rather, the focus is on the first empirical question. Specifically this dissertation examines how current manifestations of SBA reform may shape teachers’ perceptions of autonomy in the classroom.

If one accepts the hypothesis of prior research that the advent and progression of SBA reforms have altered the balance between school-site autonomy and system-wide accountability, then more work must be done to identify first, whether SBA reforms have in fact altered teachers' perceptions of their autonomy; and, second, whether that change corresponds to the underlying assumptions of the policy. For example, one might expect teachers' sense of autonomy to correlate with the implementation of increasingly more consequential forms of SBA, such as NCLB. Or, if SBA functions as it is meant to function, then teachers in schools that fail to meet performance standards should report having less autonomy than those that meet performance standards. Moreover, the more severe the sanctions that are being implemented in any year, the less autonomy teachers should have. In general, one would expect SBA to reduce a sense of autonomy for low-performing teachers but have no effect, or perhaps even a positive effect, on higher-performing teachers.

Specifically, questions remain as to whether teacher autonomy has changed over time, particularly since the implementation of No Child Left Behind Act (NCLB). I posit that teacher perceptions of autonomy have changed over time and that this change varies for teachers based on distinct school and teacher characteristics (e.g., those in low performing vs. high performing schools; and Title I schools vs. non-Title I schools). Further, I suggest that the school-level embodiments of SBA reform found in NCLB, such as adequate yearly progress (AYP) status and associated sanctions, produce differential effects on teacher autonomy that are potentially unrelated to teacher performance (e.g., tested subjects vs. non-tested subjects; and novice vs. veteran teachers). These differential effects are especially important to examine because they

could lead to unintended consequences that could undermine the policy. In the next section I present the conceptual framework for this study.

Conceptual Framework

Sociologists and political scientists who study issues of power and control in organizations broadly and teacher autonomy in schools specifically inform the current study. Specifically, I anchor this study in research by Dr. Richard Ingersoll and on notions of power and influence described by Dr. Robert Dahl (1984). Ingersoll is the preeminent scholar who uses sociological and political constructs to examine the tension between autonomy (i.e. the power to control) and accountability. Ingersoll has a depth of scholarship regarding teachers' levels of control and power in schools and classrooms and has placed particular emphasis on the tension between autonomy and accountability (Ingersoll, 1996, 2003, 2007, 2012; Ingersoll & Merrill, 2011). Specifically, Ingersoll's conceptual framework in *Who Controls Teachers' Work* (2003) and in *Teachers' Decision-Making Power and School Conflict* (1996), in which he examined the degree of power teachers and administrators held over the classroom and other school domains, guides the current study's focus on teacher-perceived autonomy in the classroom.²

The conceptual underpinnings of the study include constructs of degree, scope, and domain from Dahl's (1984) power-influence model. Specifically, this study uses Dahl's constructs of power and influence to develop a nuanced understanding of how SBA reforms affect public-school teachers' power to exercise control over their work in the classroom as they experience the tension between accountability and autonomy that

² The phrase "teacher perceptions of autonomy" appears in cited works and in the current text. While it is the most accurate phrasing to describe the phenomenon of interest in this study, it may be tiresome to the reader given the length of this study. Therefore, as others have done (Moomaw, 2005), I will use the phrases "teacher-perceived autonomy" or "teacher autonomy" interchangeably from this point forward.

SBA reform brings to education.³ Dahl defines power as the ability of one party to persuade and/or to pressure another party to perform acts which the second party is not predisposed to do. Further, he recognizes power is a relative concept. That is, power does not exist in absolute terms.

The study of power exposes who can do what to whom under what circumstances. Scholars provide some tools to recognize and depict the power of actors in given domains. In their study of state policymaking and school policy implementation, Malen and Muncey (2000) explain that to assess the autonomy of a given actor one must examine various indicators, such as “the degree of formal authority to make binding decisions”(p.201), the rules under which an actor operates, the actor’s ability to allocate resources, and “the strength of various rewards or sanctions that individual schools may anticipate or higher authorities may impose” (p. 201). Although the authors investigated school-level autonomy, researchers may adapt the same indicators to study autonomy for other levels of the public school system. For example, teachers who operate in public-school classrooms within a high-stakes SBA context may have experienced changes to their scope and degree of autonomy that either bolster or wither their ability to make “binding decisions” in the classroom.

The answers to the questions “over what?” and “over whom?” expose those who have power and what they can control. *Degree* reflects the extent to which one party may influence another in the power relationship; *scope* defines the issue or content over which an individual, group, or entity may have power; and *domain* indicates over whom or over what systems actors may exercise power. Dahl’s constructs are useful to the current study

³ From this point forward, I will either omit altogether or use sporadically the ‘public-school’ adjective because I am referring only to teachers in public schools.

because they call for interpretations of teacher autonomy that recognize gradations of control over the classroom relative to other actors and elements of the school system. Although Ingersoll does not explicitly refer to Dahl, Ingersoll's examination of issues concerning teacher power in schools includes domain (e.g., school and classroom), scope (e.g., discipline policies and hiring decisions), and degree (e.g., level of autonomy measured through qualitative and quantitative methods). Thus, it is fitting to incorporate Dahl's notions of power and influence explicitly in this study.

The general conceptual framework for this dissertation (see Fig. 1) orients the reader to the notion that power over schools and teachers flows through a system from the federal, state, and local levels of government to schools and teachers.

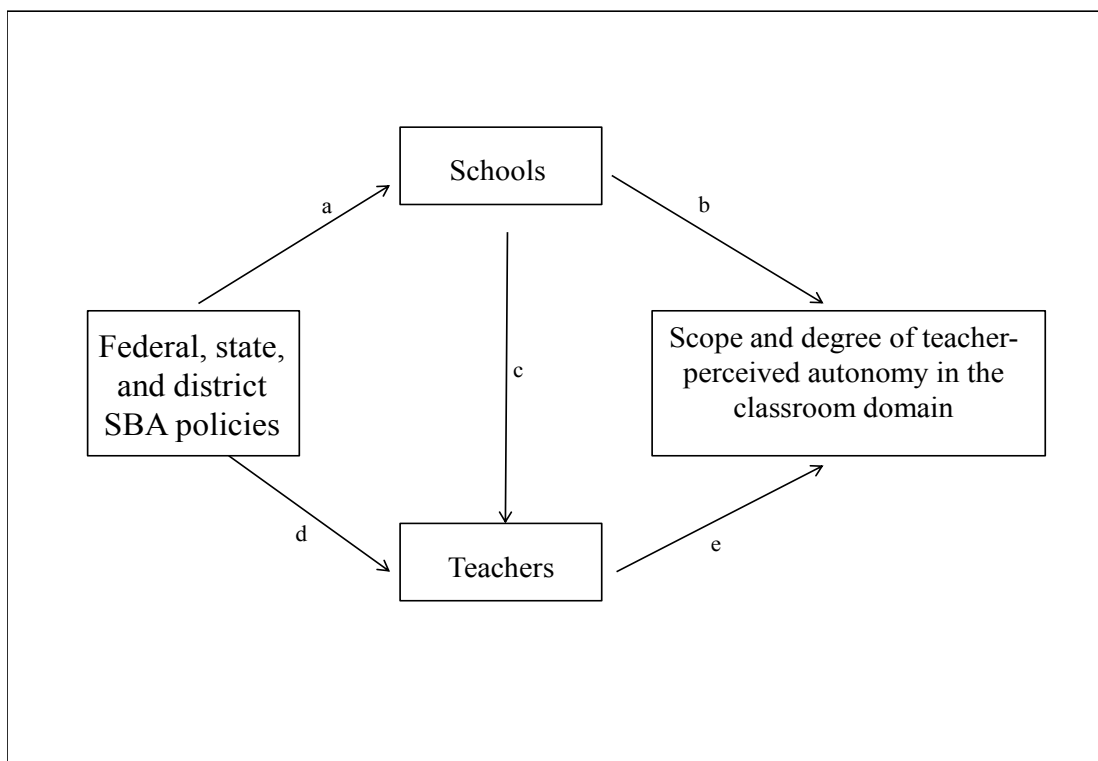


Figure 1. Relationship between standards-based reform and teacher perceptions of scope and degree of autonomy in the classroom domain.

Arrow **a** indicates that actors external to the school site (e.g. federal and state policymakers) initiate SBA reform that school-level actors must interpret and implement. For simplicity's sake, I have combined federal, state, and district actors into one box, though in reality these actors may make multiple and separate demands on schools and teachers. To illustrate, through NCLB federal policymakers decreed that Title I schools were subject to sanctions if the school failed to make AYP. However, state policymakers had a disparate response to the federal sanction policy such that in many states all schools are subject to NCLB sanctions rather than only Title I schools. This state treatment of federal sanctions is an example of both the flow of an accountability policy from external actors to school-site actors, and of the disparate demands that external actors may place on school actors through NCLB.

Arrows **b** and **c** illustrate that school-level context and actors influence the scope and degree of teacher-perceived autonomy in the classroom domain. In other words, the school mediates the effects of SBA on teachers' perceptions of autonomy – either directly through the characteristics of schools (**b**) or indirectly through specific policies and practices that influence the influence of SBA policies on teachers' work (**c**). For example, teachers in different schools may experience differing perceptions of autonomy based in part on the school context (e.g., student poverty level measured by Title I status; whether the school makes AYP); schools that are more likely to be the target of sanctions are more likely to experience threats to autonomy. Moreover, teachers in different schools may experience differing perceptions of autonomy based in part on school leadership. Principals, for example, may enact policies that either exacerbate teachers' concerns about external influences or attenuate them (Hanson, 2006). Positively or

negatively, schools may also mediate teachers' perceptions of autonomy through policies and practices, such as leadership, that influences teachers' work.

Arrow **d** illustrates that SBA policies created at the federal, state, and local levels, (e.g., curricula and testing policies of NCLB) may influence teachers. Here the figure acknowledges the possibility of a direct influence on teachers' perceptions of autonomy by external actors, one not mediated by the school. Just as specific types of schools are targeted, NCLB targets specific types of teachers (e.g., mathematics, reading and science teachers in grades three through eight). Teachers who teach tested subjects in tested grades may feel different levels of autonomy because the policy targets *their* grade and subject for testing while it does not target other grades and subjects. Arrow **e** illustrates that teachers themselves influence the scope and degree of teacher-perceived autonomy in the classroom domain based on teacher characteristics (e.g., race and experience). In other words, apart from school contexts, grade assignment or subjects taught, teachers may perceive their degree of autonomy differently based on individual characteristics.

With others, Manna (2006a) observed that "in any policy system where actors at one level of government depend on other levels to help them accomplish their objectives, a careful combination of enforcement and persuasion is required for policy success" (p. 473). The overarching hypothesis that guides this dissertation is that during the period of 1993–2008, teacher perceptions of autonomy changed and SBA reform policies enacted at the federal, state, and local levels associate with that change. To capture the nature of these changes, I incorporate Dahl's (1984) power-influence constructs. The constructs of *domain* and *scope* are consistent across all analyses. Domain always refers to the public

school classroom. Scope is always the single construct that consists of the following six measures of teacher-perceived control in the classroom domain:

1. Textbook selection
2. Teaching technique selection
3. Content, topics, and skills selection
4. Student grading and evaluation
5. Student discipline
6. Homework determination

The measure that captures the construct of teacher-perceived autonomy treats each of the six elements with equal weight; however, given the theory of action underlying SBA reform described in the first chapter, one might expect that teachers experience more change in the first three elements listed. Specifically, federal and state SBA policies emphasize standardization of course materials, content and approach to teaching. It is also possible, however, that SBA, as interpreted by school districts and schools, influence the latter three elements. Districts, for example, may require additional evaluations and testing of students, and schools may require increased discipline and additional homework assignments as part of local strategies to improve student test performance. While the connections between SBA policies and selection of materials, content, and teaching techniques might be more direct than those between grading, discipline, and homework, the latter group offer insight into policies that may be local responses to SBA reform that subsequently affect teacher autonomy. Thus, the full measure that included all six elements offers a fulsome approach to the construct of teachers' perceptions of autonomy.

In contrast to scope and domain, *degree* takes two forms in this study. I analyze the average perception of degree of teacher-perceived autonomy in the classroom domain and the variation from that average for each of four academic years of data between 1993

and 2007 (i.e., 1993-94, 1999–2000, 2003-04, and 2007-08). Second, I analyze teacher and school effects between and within schools using an indicator of whether a school failed to meet AYP status, a direct measure of NCLB, the specific manifestation of SBA that is a focus of this study in 2007–08, and a measure of level of sanctions that a school has experienced (e.g., first or second year under sanctions). I examine this AYP relationship while controlling for specific school and teacher characteristics in the full 2007–08 sample and in two subsamples of 2007–08 data based on sanction policies at the state level (i.e., whether states apply NCLB sanctions to all schools or just Title I schools).

Research Questions

Informed by scholarship that has addressed SBA reform and policy implementation, I hypothesize that the complex web of power relationships in the U.S. public school system has developed such that, for good or for ill, teacher autonomy in the classroom has narrowed since the adoption of SBA reform policies. I suggest that NCLB is the culmination of SBA reform and that whether a school makes AYP will directly influence teachers' perceptions of autonomy. I hypothesize that since the adoption of NCLB changes in teacher-perceived autonomy are neither universal nor uniform. Rather, some teachers perceive diminished autonomy in the classroom while others do not, such as teachers who have less experience compared to those who have more; teachers in Title I schools compared to those in lower-poverty schools; or teachers who teach tested subjects compared to those who do not. I further posit that notable differences in autonomy exist for teachers who work in schools that did not make AYP status compared to those who work in schools that did.

To explore my hypotheses I analyze four waves of SASS data and use the following research questions to guide the analyses:

1. How do teachers perceive their own autonomy in the classroom?
 - a. From 1993 to 2007, how has the degree of teacher-perceived autonomy in the classroom changed?
 - b. For 2007–08, what school and teacher characteristics are associated with differences in teacher-perceived autonomy?
2. How does SBA reform, represented by a school's NCLB AYP status in 2007–08, associate with teacher perceptions of autonomy in the classroom domain?
 - a. How do teachers who work in schools that made AYP perceive their autonomy compared to those who work in schools that did not make AYP?
 - b. How do teachers who work in schools in the first year of failed AYP status perceive their autonomy compared to those who work in schools in the second year or later of failed AYP status?

In sum, if SBA has permeated the classroom walls, as scholars suggest, then variations in the trend of teacher-perceived autonomy merit examination on a national level to uncover whether changes are sporadic and isolated or uniform and universal. Changes in teacher-perceived autonomy may be limited to certain types of teachers in certain types of schools, or may reflect a widespread phenomenon driven by forces external to teachers and to schools. My study addresses the breadth and depth of such changes. It also examines the extent to which the potential effects on teacher autonomy comply with the underlying assumptions of SBA – that is, do teachers in low-performing schools experience less autonomy than teachers in higher-performing schools.

Overall Design of the Study

To answer the two multipart empirical research questions listed above, I employ a two-phase research design. The study uses descriptive analyses based on SASS data collected in four academic years (1993–94, 1999–2000, 2003–04, and 2007–08) to compare changes in self-reported teacher-perceived autonomy in the classroom over time. This descriptive discussion reveals both the total variation in teacher perception of autonomy between 1993–94 and 2007–08, and the variation attributable specifically to the schools in which teachers worked. Effectively, the descriptive results offer a pattern of overall change in the average perception of autonomy as a context for the multilevel model that uses teacher and school characteristics (e.g., years of experience, subject matter taught, and school Title I and AYP status) to find specific school-level and teacher-level effects on autonomy. For the 2007–08 data, I use hierarchical linear modeling (HLM) to exploit the nested design of the NCES data. The advantage of HLM is that it more accurately reflects the relationship between school characteristics and teacher-perceived autonomy in a nested data structure compared to ordinary least squares regression (Luke, 2004). Further, HLM enables me to separate the school-level effects from teacher-level variations on perceived autonomy. This distinction is important because school-level characteristics, teacher characteristics, or, more likely, a combination of the two, may affect variation present in teachers' perceptions of autonomy.

Potential Significance

This study adds to the conceptual and empirical scholarship on teacher-perceived autonomy. Conceptually, the explicit use of Dahl's power-influence constructs to study

teachers as arbiters of SBA policies allows for a nuanced examination of teacher-perceived autonomy, which is neither static nor absolute. Empirically, the study extends scholarship through nationally generalizable results from over-time analyses of the tension between accountability and autonomy that extend to 2007–08, and from analyses that capture the effect of NCLB in the classroom (i.e., AYP status and different levels of sanctions). While I employ the same SASS variables to build the autonomy measure as did Ingersoll and May (2010; 2011) in their use of the 2003-04 and earlier data, I extend their work in four ways:

1. use of teacher-perceived autonomy as the construct of interest rather than as an indicator of another phenomenon;
2. use of Dahl's power-influence constructs to examine teacher-perceived autonomy;
3. use of SASS data from 1993–94 to 2007–08 to develop a portrait of change in teacher-perceived autonomy over time;
4. use of SASS 2007–08 to build multiple hierarchical linear models that directly measure the influence of SBA reform tools on teacher-perceived autonomy.

As evidenced by the ongoing work by Ingersoll and others, teacher-perceived autonomy in the classroom continues to be an important element of scholarship that examines a critical aspect of teachers' working conditions in classrooms and schools. However, the majority of empirical research to date on teacher autonomy does not focus on teacher autonomy as the construct of interest. Rather, researchers tend to use teacher autonomy as an indicator of other teacher-centered phenomena. Further, prior research is either not nationally representative, or does not include AYP status, which is possibly the

most public manifestation of the NCLB.⁴ In contrast, the hierarchical models of nationally representative teacher data in this study incorporate AYP status as the trigger for sanctions promoted by SBA reform, and incorporate state-level policies regarding which schools should be held accountable for AYP failure. To introduce the key elements of the relationship investigated in this study, in the next section I offer an historic overview of SBA reforms and the relationship between these reforms and teacher-perceived autonomy.

⁴ On March 9, 2011, the U.S. Secretary of Education reported to Congress that ED estimates that 82 percent of public schools will not meet AYP targets. (See: <http://www.ed.gov/news/press-releases/duncan-says-82-percent-americas-schools-could-fail-under-nclb-year>). In 2012, the Center on Education Policy updated the percentage to 49 percent. (See <http://www.cep-dc.org/displayDocument.cfm?DocumentID=403>).

Chapter 2: Literature Review

This literature review begins with a broad examination of research that has addressed the intergovernmental phenomenon of public education and the relative power over education held at each level of the public school system. The section continues with a more focused review of literature specific to the teacher level of the system: it examines research on teacher autonomy and, in turn, how autonomy interacts with teacher professionalism, job satisfaction, retention, and SBA reform.

Historical Context of SBA Reform Development and Implementation

The historical context of SBA reform is important to this study because it establishes the context of change in education policy that brought SBA reform to the fore. History is valuable because “what may seem like a relatively rapid process of reform is in fact only the final stage of a process that has been under way for an extended period” (Pierson, 2005, p. 40).

Beginning in the early 1980s, public school reformers began to incorporate requirements for standards, assessments and accountability into education policy discussions (Firestone, Fuhrman, & Kirst, 1991). This package of reforms became what is now termed *standards-based accountability reform*. According to Elmore (2002), the National Governors Association promoted the “central theory of today’s accountability reforms,” (p.32) when it introduced performance-based accountability, whereby states sought accountability for student academic performance but would leave the decision-making regarding curricula and assessment to districts and schools.

DeBray (2006) and McDermott (2003) chronicle the emergence of SBA. In the late 1980s and 1990s states began to develop and implement standards-based accountability reforms. In 1989 the Charlottesville Education Summit convened and linked standards-based reforms to the problem of school failure (Superfine, 2005). According to former Secretary of Education Rod Paige (2006), the Summit was “a landmark event in that it brought together, for the first time, nearly every state governor in one forum to talk about education” (p. 464). The Summit venue allowed governors to create a national agenda encouraging states to adopt SBA as the ideal reform. Perhaps unintentionally, the national focus and symmetry of thought at the state level laid the foundation for a federal movement that would embrace a high-stakes standards and accountability reform policy (Manna, 2006a). The Summit served as a beacon that emphasized the importance of education on the federal agenda and widened the window of opportunity for entrepreneurs that would promote a prominent federal role in education (Heise, 1994).

Calls for achievement and accountability began early in the decade and remain strong to the present day. SBA reform had a potent blend of policy tools that educators could not ignore (Malen & Muncey, 2000). For instance, SBA shifted the focus of evaluation from inputs (e.g., program expenditures) to outcomes (student assessment scores); declared student achievement as a proxy for teacher and principal effectiveness; and carried threats of school sanctions. By 1999, nearly all states had increased their scope of influence over education through the development of achievement standards and student assessments intended to align with those standards (DeBray, McDermott, &

Wohlstetter, 2005; Firestone, 2001; Fuhrman, 1999; Heise, 1994; Manna, 2006a; Sunderman, Tracey, Kim, & Orfield, 2004).

Although states implemented forms of SBA that were similar in purpose, state programs were vastly different from one another (Stecher et al., 2008). As of 1999, 14 states offered schools monetary rewards; 16 states created a system to take over failing schools; and 19 states required students to pass an exit exam to graduate high school (Education Week, 1999). Naturally, the “complexity and comprehensiveness of state accountability systems var[ied]” (Rorrer, 2004, p. 253). Thus, scholars found the comparison of accountability systems across states to be a challenge (Goertz, 2001; McDonnell, 2005). Researchers sought to evaluate state programs and cited difficulty at the state level with standards-based reform implementation that included the lack of technical expertise and the lack of implementation capacity (R. F. Elmore, Abelman, & Fuhrman, 1996; Goertz, Floden, & O'Day, 1995; Massell, 1998; Spillane & Zeuli, 1999).

As many states were building reform policies and supports, the idea of direct federal intervention was beginning to take hold but still met with resistance (DeBray et al., 2005, p. 253). At the federal level any notion of a federal government involved in national standards, core curricula, or school choice was met with legislator concern about potential breach of the tenets of federalism (Superfine, 2005). Federal intervention was already frustrating state education leaders as state leaders tried to create accountability systems and comply with existing federal programs, specifically Title I. For example, Colorado and Arizona education commissioners commented that federal involvement in schools was “confusing” and led to “hazily-defined goals” with ever-increasing regulations (Kanstoroom & Finn, 1999, p. 3). Compliance with federal programs was a

concern because during the more than 40 years of ESEA, localities nationwide had become reliant upon the steady flow of federal dollars and accustomed to the flexibility of the associated program requirements (Kantor & Lowe, 2006; Sunderman & Kim, 2007). Simply put, state education leaders were not interested in increased federal intervention though they wanted and arguably needed federal revenue.

President George H. W. Bush tried in vain to collaborate with states and pass America 2000 to further the legitimacy of SBA, offer resources to build capacity, and promote the imperative for education reform (Eisner, 1995; McDonnell, 2005). Subsequent to George H.W. Bush, President Clinton tried to pass a federal version of SBA in Goals 2000. The legislation was distinct from statutes that came before because it “embrace[d] a new approach – systemic reform” (Heise, 1994, p. 5). The Executive Branch proposed Goals 2000 as a voluntary program with the overarching goal of using grants to encourage states to develop standards, assessments, and accountability mechanisms (i.e., to embark upon systemic reform). “Together, standards, assessments, flexibility, and accountability were thought to be key components that could spur systemic reform in the American education system” (Superfine, 2005, p. 10).

Legislators and states’ rights groups met both America 2000 and Goals 2000 with arguments that these policies would lead to the “federalization of...American education policy” (Heise, 1994, p. 345), which would equate to more power over education for the federal government and less power for states and localities (Manna, 2006a). Education scholars noted that states, not the federal government, were the “key sources for new ideas” such as charter schools, and that the notion of top-down education reform was antiquated” (Kanstoroom & Finn, 1999, p. 7). Regardless of such protest, as the years

passed, SBA advocates gained momentum for what would become a hands-on federal role in the daily operation of the classroom (Rorrer, 2004)

Elements of SBA Reform Policies. SBA reform policies have many elements, including alignment between curricula and assessment, increase of teacher capacity through professional development for teachers, and use of results from standardized tests as a proxy for student achievement and practitioner performance (Hanushek & Raymond, 2002; Linn, 2000). While acknowledging the value of standards because of their potential to bring about “improve[ed] teaching and learning for all” (Firestone et al., 1991, p. 245), scholars noted early on that the SBA reform, particularly from the federal government, had several obstacles to implementation. For example, Fuhrman et al. (1991) pointed out that standards would likely measure minimum rather than aspirational achievement.

Another key element and challenge of SBA reform is the movement from input measures (e.g., per pupil expenditure) to outcome measures (e.g., promotion and graduation rates) (Cohen-Vogel, 2005; Superfine, 2005). Ostensibly, test scores would signal to educators gaps in student learning and lead to changes in instructional practices (Amrein & Berliner, 2002; Spillane & Zeuli, 1999). Policymakers contend that student performance measures both student and teacher performance. Practitioners contend that test scores are a default measure. McDonnell (2004) explains that standardized tests are the chosen instrument to measure classroom outcomes “because so few of a school’s core activities are observable and because the operational distance between state legislatures and individual classrooms is huge” (p. 9). SBA reformers assume school systems can move away from input measures to adopt standards for outcomes that will accurately measure achievement. Scholars point out that this critical assumption reveals a possible

disconnect between policy and practice in that it relies heavily on assessments to measure student achievement and educator effectiveness (Goertz, 2001; Kantor & Lowe, 2006). Nonetheless, over the years states have produced a mix of reforms that target students through testing, and teachers through certification standards (Darling-Hammond & Wise, 1985).

As SBA reform took hold in public education, local school systems experienced intense pressure to increase student performance through standardized instruction and assessment (Goodwin, Cunningham, & Eagle, 2005). While all SBA reforms attempt to align curricula, instruction, standards, and assessments with student opportunities to learn (Clune, 1993; Loucks, 2007; McDonnell, 2004; Smith, O'Day, & Cohen, 1991; Superfine, 2005), the notion that testing every student and delivering sanctions or rewards to teachers, schools, and districts can improve the quality of public education defines recent SBA reforms (Loucks, 2007; Stecher & Hamilton, 2002).

High stakes and low stakes accountability. SBA can be high- or low-stakes. Low-stakes accountability is for informational use only. That is, low-stakes accountability reform does not contain specific sanctions or rewards based on the results of the accountability instruments. Goals 2000 and early state SBAs are examples of low-stakes accountability, the former by design and the latter by default (DeBray, 2006; Fuhrman & Elmore, 1990). Goals 2000 was built on “hortatory policies that rel[ied] on either information or values to motivate action consistent with the policy’s goals” (McDonnell, 2004, p. 20). State SBAs were low-stakes primarily because states lacked the capacity to help troubled schools (Fuhrman, 1999). In contrast to low-stakes accountability, the high-stakes version holds a threat of sanctions and punitive actions for

those schools that do not comply or that underperform, and the promise of continued financial reward for those that succeed. But both low- and high-stakes SBA embrace the highly contested notion that tests are an objective measure of the intersection between classroom instruction and student achievement that can expose discrepancies in educational achievement and opportunity (Borkowski & Sneed, 2006; Clune, 1993; Darling-Hammond, 2004b; Diamond & Spillane, 2004; Dorn, 1998; Firestone, Mayrowetz, & Fairman, 1998; Linn, 2000; Mintrop, 2003; Wood, 2004).

SBA proponents agree on accountability but disagree on high- versus low-stakes policies (Abrams, Pedulla, & Madaus, 2003). Firestone, Mayrowetz, and Fairman (1998) describe in broad terms the two sides of the SBA high-stakes versus low-stakes issue. High-stakes opponents claim that standards-based accountability shrinks the school curriculum, reduces the professionalism of teaching by “deskilling” teachers, and propagates anxiety among both students and teachers who face sanctions for outcomes that may or may not be under their direct control. High-stakes proponents claim that only measured results are valid and that assessment ensures compliance with a “productive and effective system” (Firestone et al., 1998, p. 95). The two viewpoints do not differ in their view that assessment-driven policies can be a “powerful lever for shaping instruction” (Firestone et al., 1998, p. 95); but they disagree on the shape accountability should take to be most effective in improving student achievement.

The growth of the federal role in education. By 1999, nearly all states had developed student achievement standards and were working on assessments aligned with those standards (DeBray et al., 2005; Firestone, 2001; Fuhrman, 1999; Heise, 1994; Manna, 2006a; Sunderman et al., 2004). As accountability reform took hold in more

states across the nation, policymakers at the federal level recognized an opportunity to leverage state actions to embolden federal influence over schools and standardize the uneven state-level implementation (Goertz, 2005; McDermott & Jensen, 2005; Shipps, 2011). As mentioned briefly above, President Clinton championed a federal version of low-stakes SBA in Goals 2000. The legislation was distinct from statutes that came before because it “embrace[d] a new approach—systemic reform” (Heise, 1994, p. 5), as opposed to prior federal initiatives that were targeted to specific populations of students or specific types of schools.

The Clinton Administration proposed Goals 2000 as a voluntary program with the overarching goal of using grants to encourage states to develop standards, assessments, and accountability mechanisms (i.e., to embark upon systemic reform). “Together, standards, assessments, flexibility, and accountability were thought to be key components that could spur systemic reform in the American education system” (Superfine, 2005, p. 10). Goals 2000 was met with arguments that these policies would lead to the “federalization of . . . American education policy” (Heise, 1994, p. 345), which would mean the federal government would have more power and states and localities would have less power over public schools (Manna, 2006a). However, as the years passed, SBA advocates gained support at the federal level and Congress passed legislation that would lead to a hands-on federal role in the daily operation of the classroom (Rorrer, 2004).

Kingdon (1995) observed that “a change of administration is probably the most obvious window [of opportunity]” (p. 168). The new administration of 2000 led to a sea change in the federal role in public education. In 2001, SBA policy entrepreneurs successfully reframed their policy solution as a federal initiative rather than a state or

national movement. They took advantage of the new Bush administration as the triggering event that would lead to SBA adoption at the federal level (DeBray, 2005). In 2001, concerns about states' rights and federalism gave way to political maneuvering needed to pass NCLB (Rudalevige, 2003). A Republican-majority Congress (many of whom had only six years previously voted to abolish the United States Department of Education) passed NCLB and applauded its call for accountability. The administration not only successfully shifted the focus of the largest federal education program from inputs to outcomes but also upset the balance of power that had historically served public education (Wong & Sunderman, 2007). Effectively, the legislative and executive branches turned away from longstanding notions of federalism and toward systemic reform sponsored and embraced at the federal level (DeBray, 2005; Superfine, 2005).

Scholars describe the law as a sea change that shifts the distribution of power over classroom practices and student achievement away from teachers and toward other actors and other levels of the education system, specifically toward the federal level (Hannaway & Woodroffe, 2003; Hollingsworth et al., 2007; Karen, 2005; Sunderman & Orfield, 2006). The passage of NCLB equated to a high-stakes SBA policy by the federal government (Brown & Cooper, 2011; Guggino & Brint, 2010) and marked the beginning of federal high-stakes accountability education reform. The Act implicates faculty, administrators, and state regulatory agencies in the failure or success of students and contains specific punitive actions for "failing" schools (Sunderman & Kim, 2007). Simply put, SBA reforms instituted classroom-level scrutiny as the basis for school system success or failure (O'Day, 2002). While NCLB maintains the historical ESEA focus on Title I schools and students, controversy surrounds the motivation for and

mechanics of the law. Some have viewed the Act as a new phase in a national commitment to public education; others have perceived NCLB as laudable but unrealistic and unachievable; but still others have identified the law as a threat to federalism (Umpstead, 2008; Wanker & Christie, 2005). For example, Manna (2006a) posits that the intent of NCLB may be to increase federal control over education through accountability that is operationalized as student testing:

The logic of standards and test-based accountability in education suggests a compelling logic of control. Why have standards and tests at all? One reason is so that leaders at upper levels of the policy system can compel actors at lower levels to produce desired results. Lacking confidence in state governments to guarantee high student achievement, federal officials have attempted to seize more control with NCLB's prescriptive accountability components. (p. 474)

In other words, integral to the NCLB brand of high-stakes SBA reform is the federal mandate of annual standardized state testing of core academic subjects (e.g., English language arts, mathematics, and science) that would compel state and local compliance with standards across the district, state, and nation and would delegate accountability to the federal level of the system. The cornerstone of NCLB is the measurement of student achievement, along with standards, assessments, and accountability (Hamilton et al., 2007). Thus, NCLB flows through the entire public school system: the federal level prescribes sanctions; the state level defines assessments and reports results; and the district, school, and classroom levels implement curricula and the testing of students.

The NCLB test-centered approach is not the only manner of available assessment. For example, educators can use student portfolios, locally developed assessments, and/or cycles of testing rather than annual cumulative tests to achieve accountability and measure progress towards standards (Elmore, 2002). In theory, high-stakes accountability

is an effective way to achieve reform in schools because assessments can expose discrepancies in educational achievement and opportunity and induce improvement (Borkowski & Sneed, 2006; Clune, 1993; Darling-Hammond, 2004b; Diamond & Spillane, 2004; Dorn, 1998; Firestone et al., 1998; Mintrop, 2003; Wood, 2004). In practice, testing may serve as the most effective, if not the only, link between the passage of the federal policy and the ability of the federal government to enforce implementation because test results are measureable and observable which enables them to span the distance between the federal level and the classroom.

Notwithstanding the ongoing efforts by states to seek waivers from ED on aspects of NCLB, as passed, the most potent measure of school performance in the NCLB legislation is Adequate Yearly Progress (AYP).⁵ The AYP measure is the critical indicator of a school's movement toward or away from academic goals. Based on NCLB requirements, the state education agencies (SEA) define the measure and the federal education agency approves or rejects the definition. Originally, NCLB requires that all students be proficient in language arts and mathematics by 2014 (Manna, 2006a). The law requires tests in reading and mathematics annually in grades three through eight and once again in high school. Further, the schools must test students at least once in science between grades three and five, six and eight, and nine and 12. Notwithstanding the current deadlines and sanctions, states continue to lobby the Department of Education (ED) for waivers and revisions of the Act. When awarded, ED is effectively exchanging strict adherence with the Act for state-level commitment to the Act, without removing the

⁵ Discussed in more detail in Chapter 5. At the writing of this dissertation states continue to apply for waivers that change the requirements and timelines of AYP. As for 2012, the Obama Administration has signaled its willingness to approve waivers and accommodate state plans that alter AYP requirements as written in the 2001 NCLB Act.

motivation for Title I schools to focus on student outcomes as measured by standardized test results (Wanker & Christie, 2005).

My dissertation embraces the definition of SBA reform as systemic alignment to bring about measurable improvement in student achievement, and postulates that NCLB is the quintessential high-stakes SBA reform policy. Two NCLB provisions, AYP status and AYP sanctions, are central to the current study. Again, NCLB requires each state to establish a definition of AYP that serves as a measure of district and school-level academic achievement and to report on those measures for each student subgroup rather than for the “average student.”⁶ Those districts and schools in which students meet proficiency levels make AYP (i.e., pass); those in which students do not meet proficiency levels fail to make AYP (i.e., fail).

Although all schools are subject to AYP reporting under NCLB, the federal law stipulates sanctions only for Title I schools that fail AYP. ED delineates specific actions that state, district, and school actors must take when schools fail to make AYP for two consecutive years. Schools do not suffer sanctions following the first year of failure but must improve achievement scores by year two. For those schools that do fail to make AYP for two consecutive years, NCLB begins Year 1 of school improvement sanctions, which include school choice and supplemental education services (e.g., tutoring). ED estimated that in the 2003-04 academic year 3.9 million public-school students were eligible to transfer to a non-failing school in the district or to another district if all schools in the home district had failed to meet AYP (Bathon & Spradlin, 2007). Further, after three consecutive years of failed AYP status, NCLB sanctions required more than 2,800 schools to provide supplemental education services (GAO, 2006). By 2008, 411 school

⁶ Previous versions of ESEA required only school-wide averages on assessments.

districts in 27 states had entered some level of NCLB sanction phase. In that year, California had the largest number of failed schools with schools in 97 of its districts, and 47 districts in Kentucky failed based on school AYP status (Williams, 2008). By 2009, Oregon had reported that of the 500 Title I schools (which includes over 300 elementary schools), more than 100 faced the school choice and SES sanctions (Hammond, 2009).

Again, NCLB prescribes sanctions only for schools that receive Title I funding. Some scholars argue that a single AYP measure for all students in a system where failure of any category of students (e.g., by race, ELL, or special needs status) results in a failed school and district puts a particular burden on those schools which serve primarily lower-income and minority students (Marx & Harris, 2006). Thus, perhaps due to a confluence of challenges from school context, the penalties of high-stakes in NCLB will most likely fall on the schools and teachers that face the most challenges. Furthermore, schools and teachers experience additional federal pressure from the Act's stated sanctions and from its ability to create uncertainty for education professionals who attempt to implement accountability measures in schools, districts, and states with varying capacity to enact change (Sunderman & Orfield, 2006).

At the school level, high-stakes accountability policies, like NCLB, have the potential to further widen, rather than remedy, a bifurcated public school system based on student performance. For example, a 2004 case study of four Chicago schools that included two low-performing and two high-performing schools found that although school staff in all four schools were more aware of and focused on the assessments, those in the low-performing schools were driven to "superficial responses based on threats" (Diamond & Spillane, 2004, p. 1157). In contrast, the authors found that the teachers in

the high-performing schools that did not operate under the threat of sanctions felt driven to make changes that brought about continued and increased student achievement. The authors suggested that low-performing schools may experience a more frenetic or “Christmas tree” (Diamond & Spillane, 2004, p. 1159) pattern of reactions to NCLB that could result in a focus on probationary status and sanctions that influence teacher autonomy in the classroom differently based on the subject matter taught, the grade level taught, and school characteristics.

SBA Implementation across the public school system

To situate “coordination and control” of SBA reforms across the federal, state, district and school levels of the public education system, the four subsections that follow focus on reform activities developed at each respective level of the system.

Federal level. Elmore and McLaughlin (1982) describe the intergovernmental relationship perpetuated by ESEA as a “federal-state-local division of labor...tight enough to establish a federal presence in local school systems but loose enough to allow for the play of state and local interests” (p. 163). Generally speaking, early federal policies promoted federal priorities through hands-off funding mechanisms—formulas and outright grants. The federal government declared program goals and provided funding; the state government elucidated program requirements and transferred funds; the localities developed program content and delivered services. To illustrate, during the Johnson Administration’s effort to bring about equity in public education, under the “Great Society” program, Congress and the Administration inaugurated ESEA of 1965. Elmore and McLaughlin (1982) explain that the federal government’s goals in 1965 were threefold: to supplement resources for targeted student populations (e.g., low-income); to

build capacity at the state and local levels that would allow for successful long-term implementation of ESEA programs; and to promote the federal agenda for a better society.

Since 1965, the federal government has followed an ebb-and-flow pattern of influence and control over local education processes. Beginning with the 1980s and *A Nation at Risk*, to the 1990s with Goals 2000, to the 2000s with NCLB, some scholars believe that the federal government has shifted away from primarily an agenda-setting role and in addition to that role has adopted a direct school-level role that has culminated in standards, accountability, and NCLB (Campbell, R. 1985; Hill, 1999; Hill, Cross, & Kilgore, 2000; Kanstoroom & Finn, 1999; Lieberman, 1982; Manna, 2002; McDonnell, 2005). Scholars document that since the 2001 Act, school staff have experienced the federal role in education more directly in four ways: first, requirements of “scientifically-based” curricula and instruction; second, definition of core academic subjects; third, requirements for teacher certification; and, fourth, sanctions associated with the failure to meet AYP status (Mintrop & Sunderman, 2009). Table 1 lists the NCLB-prescribed sanctions associated with each consecutive year of failure.

Table 1

AYP Sanction

Year	Label	Sanction
1	School improvement	<ul style="list-style-type: none"> - Technical assistance by the state and/or district to include assistance in data analysis, professional development, and resource allocation - Public school choice: each student is eligible to transfer to another public school in the same district - Development of a two-year school improvement plan “in consultation with parents, school staff, the local educational agency, and other experts, for approval by the LEA. The plan must incorporate research-based strategies, a 10 % set-aside of Title I funds for professional development, extended learning time as appropriate (including school day or year), strategies to promote effective parental involvement and mentoring for new teachers.”
2	School improvement	<ul style="list-style-type: none"> - All of the above plus: - LEA to provide supplemental educational services to students from low-income families (i.e., free extra academic help, such as tutoring or remedial help, provided to students in subjects such as reading, language arts, and math. This help can be provided before or after school, on weekends, or in the summer)
3	Corrective action	<ul style="list-style-type: none"> - All of the above plus: - One of the following: <ul style="list-style-type: none"> • Replace school staff responsible for the continued failure to make AYP; • Implement a new curriculum and professional development based on scientifically-based research; • Significantly decrease management authority at the school level; • Extend the school day or school year; • Appoint an outside expert to advise the school on its progress toward making AYP in accordance with its school plan; OR • Reorganize the school internally.
4	Restructuring	<ul style="list-style-type: none"> - Continue to offer public school choice and supplemental services - LEA must prepare a plan to carry out one of the following: <ul style="list-style-type: none"> • Reopen school as charter school • Replace principal and staff. • Contract for private management company of demonstrated effectiveness • State takeover • Any other major restructuring of school governance
5	Implementation of Restructuring	No later than first day of school year following Year 4 of AYP sanctions, the school must implement the school restructuring plan selected from the list above.

Note. Adapted from <http://www2.ed.gov/policy/elsec/guid/secletter/020724.html#chart>

Wanker and Christie (2005) point out that NCLB has two central elements that distinguish the law from prior iterations of ESEA:

First, it represents a more systemic approach to achieving reform and improvement, tying together a variety of requirements and incentives in areas ranging from student testing, school safety, and reading instruction to professional development for teachers and technical assistance for low-performing schools. Second, it significantly raises the stakes—for states, districts, and schools—for failure to make steady, demonstrable progress toward improving student achievement (p. 58).

In concurrence with many scholars who identify NCLB as the farthest reaching federal education policy to date (Borkowski & Sneed, 2006; Goertz, 2001; Kantor & Lowe, 2006; Manna, 2006a; McDermott, 2003; McDonnell, 2005; Meier & Wood, 2004; Rorrer, 2004; Sunderman & Kim, 2007; Sunderman & Orfield, 2006; Superfine, 2005), this study place NCLB at the apex of high-stakes SBA reform policies in schools. The distinction of NCLB as the epitome of SBA reform is important because I hypothesize that key aspects of the law, specifically adequate yearly progress (AYP) and associated sanctions, have direct associations with teacher-perceived autonomy in the classroom.

NCLB's adoption of testing as the singular indicator of academic achievement "reflect a gradual but steady shift from the use of tests as measurement instruments designed to produce information to a reliance on tests to influence policy and instruction" (Hamilton, 2003, p. 28). In contrast to low-stakes accountability, the theory of action behind high-stakes accountability is that a threat of sanctions and punitive actions for those schools that do not comply or that underperform will spur improvement (Elmore, 2002). Debate exists about the effectiveness of sanctions (Mintrop & Sunderman, 2009). Proponents of high-stakes accountability posit, among other things, that the high stakes will motivate students to do their best; will motivate adoption of quality curricula; will

improve teacher quality; and will enable administrators to implement targeted professional development for teachers (Amrein & Berliner, 2002; Hamilton, 2003).

Opponents of high-stakes accountability argue that implementation will increase challenges to staff and students of schools that already face difficulties (e.g., high-poverty, high minority and ELL enrollment, and high special education enrollment) and will create instability in schools that already face higher-than-average teacher turnover rates (Mintrop & Sunderman, 2009; Olsen, 2008; Santoro, 2011). Opponents also contend that high-stakes reforms will demoralize teachers because of the threat of sanctions associated with poor results from standardized tests, and because the curricular and pedagogical changes will focus on test content rather than knowledge content based on practitioner expertise (Hamilton, 2003).

Prior to NCLB, some states, like Florida and New York, implemented sanctions with mixed success. For example, Goldhaber and Hannaway (2004) studied five schools in Florida to identify the impact of Florida's A+ Accountability Plan, implemented in 1999, on three failing schools and two high-performing schools. The overarching mechanism of the A+ plan was to introduce competition through vouchers. Students at failing schools would be able to use their vouchers to go to another school including private and religious schools in the state. This market-driven approach is not explicit in NCLB but does resemble the transfer policy in NCLB that applies during the second consecutive year of AYP failure. A second element of the A+ plan was public awareness of school performance, a sanction also in NCLB, such that those schools that earned an "F" based on student performance on the state assessment would carry the "failing" label. The authors found that in two years after A+ implementation, the high-performing

schools retained their “A” grade, while one failing school achieved a “C”, a second achieved a “D”, and the third remained failing.

Although this exploratory study does not permit generalization to a wider population, it can suggest possible outcomes for other teachers, schools and districts. Specifically, the authors noted that teachers and principals in the “A” schools felt tremendous pressure to retain their grade. Rather than continue doing what had originally earned their “A” status, they narrowed the curriculum and spent time in class on practice exams. One interviewee commented, “We are an A school and we hate it” (p. 601). Conversely, the staff in the failing schools, with student populations marked by higher poverty, more ELL students, and more IEP students, valued the A+ program for bringing higher expectations to students that educators and policy makers customarily assumed to be low-performers.

In addition to punitive measures (i.e., sanctions), the federal power structure has programmatic elements tied to NCLB (Santoro, 2011). For example, ED describes the Reading First grants program as one that “focuses on putting proven methods of early reading instruction in classrooms.”⁷ The funding mechanism operates similarly to other federal programs. That is, states apply for grants, and if awarded, the federal governments send resources to the states that in turn send resources to the localities. “The NCLB Act federalized reading policy, and, along the way, debates about how best to teach reading moved from teacher lounges and school boards meetings to state capitols and finally to the White House and Congress” (Allington, 2006, p. 3). Scholars have cited schools in which Reading First program requirements have led to increased instructional time in phonemic awareness, phonics, vocabulary, fluency, and comprehension (Gamse, Jacob,

⁷ <http://www2.ed.gov/programs/readingfirst/index.html>

Horst, Boulay, & Unlu, 2008) because to secure federal funding, the program prescribes detailed instruction techniques in addition to programmatic goals of instruction.

Notwithstanding the flow of funds, Reading First is a departure from other ESEA programs because unlike prior federal initiatives, Reading First mandates specific instructional practices within the classroom rather than encouraging general program goals while leaving details on implementation to district and school staffs (McDermott & Jensen, 2005).

State level. Just as intervention at the federal level must span multiple layers of the education system to reach the classroom, so do state policies work their way from the capitol to the classroom. Successful implementation of state policies depends “on specified procedures for implementing and monitoring policy intentions” (Darling-Hammond & Wise, 1985, p. 315). Under earlier versions of ESEA, the state role was that of “monitoring, evaluating. . .and accounting” (Elmore & McLaughlin, 1982, p. 163) for federal funds. During NCLB implementation, states needed to develop new controls to comply with federal policy. For example, states developed certification requirements for a candidate to be a Highly Qualified Teacher (HQT), defined AYP benchmarks, and created (or outsourced) assessments aligned with curricula (Goertz, 2005; Wanker & Christie, 2005).

McDermott and Jensen (2005) described shortfalls in SEA capacity to implement NCLB. States would need the human and financial capital to support localities’ efforts to deliver standardized instruction and measure achievement, and would need data management systems to meet the reporting requirements of the Act. In theory, under NCLB states’ scope of influence over education widened and their degree of power over

districts and schools intensified. Practically, though, as states began to implement NCLB, state-level gaps in professional expertise with curriculum and instruction diminished states' abilities to implement NCLB requirements and influence teachers in the classroom (Goertz, 2005; Sunderman & Orfield, 2006). For example, in 2007, Birman, LeFloch, Klekotka, Ludwig, Taylor et al. found state-to-state differences in HQT definitions. The study found that all states met the NCLB requirement that teachers have a bachelor's degree and full certification in the state. However, the component of HQT that requires states to determine how teachers would "demonstrate adequate content knowledge in each subject taught" (p. 12) revealed state-level variation in the definitions of new and veteran teachers, the use of the Praxis exams as indicators of adequate knowledge, and the qualifications for veteran teachers to meet the High Objective Uniform State Standard of Evaluation.

Perhaps due to issues of capacity or due to a desire to keep the federal law at bay, states have actively sought ED waivers regarding compliance with certain NCLB requirements (Sunderman, 2006). A 2006 report from the Council of Chief State School Officers (CCSSO) documented the number and types of waivers states requested and how ED responded to those requests. The study found that "between 2003 and 2005, nearly every state submitted at least one request to amend its [accountability] plan," (p. 1). Specific to 2005–06, the authors noted that

States asked for changes related to their academic achievement standards, linking of achievement standards from old tests to new ones, the use of college admissions or other exams in lieu of State tests at the high school level, delays in using the results of newly assessed grades in AYP determinations for up to three years, recalculation of AYP using results from subsequent administrations, banking of test results, changes in the attribution of scores due to shifts in a testing window, and the use of formative assessment results for AYP purposes. (p.11)

In other words, states sought to manage the law such that they would achieve some level of compliance but would “minimize the law's encroachment” (Edelman, 1992, p. 1535). The CSSSO study concluded that overall ED met state requests for waivers with flat denial or uneven response in 2006 (Forte & Erpenbach, 2006). For example, North Dakota and Minnesota made very similar requests to extend development of AYP calculations. ED approved Minnesota’s request but denied North Dakota’s. Similarly, ED approved use of growth models for North Carolina and Tennessee but denied Florida.

More recently, the Obama Administration has engaged in state waiver transactions. Just as the advent of the Bush presidency in 2000 opened a policy window for education policymakers, the Obama presidency opened a new window in 2008. In August 2011, President Obama announced that in exchange for flexibility with “cornerstone requirements of the No Child Left Behind Act, including the 2014 deadline for all students to be proficient in math and reading/language arts” (McNeil & Klein, 2011, p. 11), states would commit to promoting college and career readiness for students and would agree to target their efforts toward improvement of academic achievement in the lowest 15 percent of schools. The Obama Administration also granted the flexibility for the local level to design its own “intervention [i.e., sanctions] for failing schools” (p. 20).

The Center on Education Policy (CEP) reports that as of November 2011, 39 states plus Puerto Rico and Washington D.C. have been engaged in the waiver application process, and as of February, 2012, 11 states have received waivers with the remaining 27 awaiting ED’s response. CEP noted that although states do include plans to report student achievement measures by subgroups (although not necessarily consistent

with NCLB subgroups), none of the states plan to continue the school-choice or supplemental education services provisions that are the NCLB sanctions for two and three consecutive years of failed AYP status, respectively (Center on Education Policy, 2012). While states continue to seek negotiated compliance through waivers, based on the state-specific HQT definition they also continue to exert influence over policies affecting teachers, such as deciding who can teach, what they teach, and how they teach. Further, the majority of states have extended the reach of AYP sanctions from only those schools that receive Title I funding to any school in the state that fails AYP status. As of the 2009–10 academic year, NCES reported that 32 states applied sanctions to all schools that failed AYP status.

District level. Historically, school districts have epitomized local control of public education (Campbell, Cunningham, Nystrand, Usdan, 1990). According to Meier and Wood (2004), in 1930 about 200,000 local school boards existed in the United States. In contrast, in 2004 “with twice as many citizens and three times as many students in our public schools, we ha[d] only 15,000 [districts]. . .once one of every 500 citizens sat on a school board; today it's one out of nearly 20,000” (p. 66). Scholars have noted that school boards are “less vital policy actors than they were in decades before 1980. . .having been replaced by policy elites” (Shipps, 2011, p. 259). Just as the change in value of a company’s stock can signal strategic or organizational change in the firm, changes in school districts can be a bellwether for shifts in the pattern of power over levels of the public education system (Campbell et al., 1990). For example, under IASA (the version of ESEA just prior to NCLB) districts had the power to design corrective action programs for schools in which students did not meet proficiency goals. Under NCLB, districts have

lost some of this scope of power because the federal policy dictates the corrective actions (Goertz, 2005).

Nonetheless, districts are important to the successful implementation of SBA reforms. While states adopt standardized curricula, districts are the link between adoption and implementation. District staff may deflect and defer reform through symbolic gestures or they may assist and enable implementation through substantive change (Ogawa, Haymore Sandholz, Martinez-Flores, & Paredes Scribner, 2003). Literature on the district role in SBA reform is most often technical and attends to professional development for teachers or creation of standardized materials for classrooms that reflect state standards.⁸ In fact, Goertz (2001) points out that districts are often overlooked in the SBA discussion because under SBA states have the responsibility to take over schools and do not need assistance or permission from districts to do so. In a longitudinal study that investigated the relationship between school district implementation of high-stakes accountability reform and its effect on student outcomes, Opfer, Henry and Mashburn (2008) observed that “responses to accountability mandates are generally believed to occur at three levels:

1. changes in teachers’ behaviors directed toward improving their teaching,
2. changes in school level support for improving student outcomes, or
3. changes in school district support for improving student outcomes” (2008, p. 301).

The research team examined school districts in six southern states: Georgia, Kentucky, Mississippi, North Carolina, South Carolina, and Tennessee. The team developed “a

⁸ See for example, Opfer, Henry, & Mashburn, 2008 and Darling-Hammond, 2004

straightforward line of reasoning” to anticipate the school district theory of action surrounding high-stakes SBA implementation:

Districts are compelled to develop coherent instructional policies, including professional development opportunities; these policies focus teacher attention on instructional improvement; and this instructional focus results in improved teaching and learning. (Opfer et al., 2008, p. 301)

That is, the researchers recognized a pattern of school district involvement and SBA reforms also documented by McLaughlin and Talbert (2003) in which districts that develop highly evolved and systemic focus on instructional improvement, professional development, and data-driven accountability become more involved in teaching and student achievement. In other words, a district with technical capabilities would have greater opportunity to control schools that operate under NCLB.

School level. SBA reforms have brought terms such as “reconstituted,” “low-performing,” “failing,” and “in need of improvement” to schools. Historically, schools have been symbols that reflect the status of the neighborhood, the achievement level of the local population, and the quality of the resident teaching force (Airasian, 1988). Under SBA, schools face challenges not only to their programmatic offerings but also to their symbolic meaning to the community and to those who work in schools. Although a variety of personnel operate schools, this section discusses school leaders, that is, principals.

The role of the principal has evolved in tandem with the role of schools as the primary providers of education. Goodwin, Cunningham, and Eagle (2005) offer a historical portrayal of the principalship. During the 1950s principals were central to the school-wide management of external demands related to desegregation and curricular changes following Sputnik; and internal demands such as staffing teachers and

developing discipline policies. During the 1960s and 1970s, principals became arbiters of teacher union contracts, enforcers of civil rights legislation, and overseers of federal programs (Davis, Darling-Hammond, LaPointe, & Meyerson, 2005). During the 1980s, however, principals' attention turned toward increasing student achievement and managing the stated goals of site-based management to include teachers, parents, and community leaders in school policies.

Firestone and Wilson (as cited in Hanson, 2006) describe how teachers used to be selective and specific about what they wanted from principals. They wanted principals to provide order in the school and protection from parents and community members so that they, the teachers, would be able to have “autonomy to teach the way they want[ed] and often what they want[ed]” (p. 95). As teachers protected their classroom from outsiders, principals focused on school-wide program management. Some teachers operated in isolation, while others turned to their peers for feedback on how to organize and deliver instructions.

Through the 1980s and early 1990s, principals opted not to breach the classroom door because they did not want to risk disturbing the peaceful coexistence with teachers (Hanson, 2006). However, during the mid-1990s, school-level accountability for student achievement brought new pressures and responsibilities to the principal (Kafka, 2009). Ingersoll (2003) points out that this power differential between teachers and principals comes with increased pressure for principals to increase schoolwide academic performance. Other scholars found that district leaders pressure principals in areas of student performance and hold them accountable for low achievement while at the same time limit their power over schools (Lyons & Algozzine, 2006).

Diamond and Spillane (2004) document how principals have broken the informal agreements under which teachers and principals had operated prior to SBA. Noted earlier, the team conducted a study of probationary schools in Chicago and found that principals in those schools used the accountability policy to legitimize their own role as school leaders and to push reforms that teacher may otherwise have resisted. The researchers stated that in one school,

“Dr. Smith [principal] used the accountability policy to get teachers’ attention and motivate them to change. Her first year at the school, she used accountability policy as a way to legitimize her push for teachers to improve their practice. She reminded teachers that she could and would replace those who were not pulling their weight” (p. 1156).

The case study indicated that in the two schools that failed to make AYP, principals had greater need to focus teacher attention on student achievement and had greater ability to do so because of the sanctions associated with AYP status.

Conversely, in high-performing schools, Diamond and Spillane (2004) found that school leaders used rewards more than sanctions to motivate teachers. For example, principals would incorporate praise for teachers whose students excelled on schoolwide assessments into professional development programs and would publicize the student performance to the community at large. Principals in the two high-performing schools tried to energize teachers to maintain high performance through comparisons to other schools in the community.

The authors concluded that in all four schools, principals were able to focus teacher attention on student achievement because of the performance requirements in NCLB. Nonetheless, the scope of principal influence appears narrowed due to the trickle-down nature of the federal agenda for accountability. For example, the NCLB

requirement for a highly qualified teacher has led some principals to perceive that they no longer have power over hiring decisions at the school (Rice, Roellke, Sparks, & Kolbe, 2006). Although principals were able to push reform in Chicago, ultimately it was not their reform — they were agents of the federal, state, and district policies.

Research on Teacher Autonomy

Pearson and Moomaw (2006) stated “autonomy is difficult to isolate” (p. 44) because of the confluence of autonomy with other constructs, such as teacher job satisfaction, teacher professionalism, and teacher motivation. Research shows that teachers have an array of preferences in the workplace that statistically associates with autonomy and that these preferences for characteristics of the workplace vary across schools (Clotfelter, Ladd, & Vigdor, 2006). Principal among the non-pecuniary workplace dimensions are school leadership, class size, school climate, school location, student demographics, and parental behavior (Beteille, Kalogrides, & Loeb, 2009; Boyd et al., 2011; Rice et al., 2006). Academic achievement and accountability measures, such as AYP status, also mediate teacher preferences regarding which schools they would select as their workplace and have the potential to affect teacher perception of autonomy (Feng, Figlio, & Sass, 2010).

Although in this study I explore teacher-perceived autonomy as a construct distinct from other workplace conditions, scholars tend to study teacher autonomy as a function of, or relative to, other conditions. For example, Pearson and Moomaw (2006) described teacher autonomy as “a common link that appears when examining teacher motivation, job satisfaction, stress (burnout), professionalism, and empowerment” (2006, p. 44). To try and disentangle the findings of prior research that used teacher autonomy as

an indicator of another phenomenon, I present literature in three sub-sections: teacher autonomy and professionalism; teacher autonomy and job satisfaction; teacher autonomy and retention. I then present a fourth subsection focused on research that has examined the relationship between teacher autonomy and SBA reform.

Teacher autonomy and professionalism. Sociologists reference a “professional model” for occupations that typically includes the following three elements: technical knowledge gained through education and training, service ethic toward client, and an occupation-wide enforcement of standards (Burbules & Densmore, 1991a; Downie, 1990; Guggino & Brint, 2010; Sykes, 1999). Ingersoll (2003) adds that “one of the most important criteria distinguishing the degree of professionalization and the status of an occupation is the degree of power and control practitioners hold over workplace decisions” (pg. 224). Teachers have sought professional status to legitimize their authority as educators; however, the lack of control over higher-level organizational resources, such as funding, curriculum choice, and discipline policies, impedes teacher claim on professionalism (Conley, 1991; Ingersoll, 2003).

Thus, it should be no surprise that the professional status of teachers has been a longstanding debate among scholars, including historians, philosophers, sociologists, and organizational theorists (Burbules & Densmore, 1991b; Firestone & Bader, 1991; Palmer, 1953; Strike, 1990; Sykes, 1999; Talbert & McLaughlin, 1994). Ingersoll and Merrill (2011) delineated three aspects of the polemic surrounding teachers as professionals:

To some, the essence of a profession is advanced training and, hence, the way to best professionalize teaching is to upgrade teachers’ knowledge and skills through professional development. For others, the essence of a profession lies in the attitudes individual practitioners hold toward their work. In this view the best way to professionalize teaching is to instill an ethos of public service and high standards—a sense of professionalism—among teachers. For even others, the focus is on the organizational conditions under which teachers work; in this view,

the best way to professionalize teaching is to improve teachers' working conditions. (p. 185)

These three positions illustrate that research presents differing views of whether and how teaching could be declared a profession, depending on the goal of the professional status – to signal technical expertise, to model a value system, or to create appealing work environments.

A concept related to *professional* is *professionalism*. Over the last few decades, scholars have documented a shift in teaching away from the notion of a profession and toward a notion of professionalism. Based on the theoretical foundation of teachers as professionals and teacher professionalism, scholars have conducted empirical studies to try to capture the relationship between teacher autonomy and professionalism. For example, Pearson and Moomaw (2005) used a Teacher Autonomy Scale (TAS) that they had developed in the 1990s (described in more detail in the teacher autonomy and SBA reform subsection) to investigate the hypothesis that “autonomous teachers would demonstrate less on-the-job stress, greater work satisfaction, perceived empowerment, and a high degree of professionalism” (p. 42). The researchers considered curriculum autonomy (items related to the selection of class materials and instructional planning) and general teaching autonomy (items related to classroom standards of conduct and decision-making ability) as distinct measures. Based on 171 respondents who worked in primary and secondary schools across three school districts in Florida, the authors used multivariate analysis of variance to determine that higher levels of curricular autonomy are associated with lower levels of on-the-job stress, and that as general teaching autonomy increased so did teachers' perceptions of empowerment and professionalism.

Effectively, teachers who perceived themselves as autonomous also perceived themselves as professionals.

More recently, in spring 2007, Guggino and Brint (2010) surveyed over 700 California teachers. This study highlights three points relevant to the current study. First, autonomy and professionalism are related constructs in teaching. Second, though related, the two constructs are distinct and one can investigate them as such. Third, teachers value autonomy and perceive themselves to be professionals. The goal of the study was to “determine how highly accomplished teachers perceived the No Child Left Behind Act’s impact on their practices in the classroom and their status as professionals” (p. 3). The researchers focused on National Board Certified teachers because they perceived these teachers to be the most highly skilled, and therefore, the best representatives of teachers as professionals. The authors stated that overall, 84 percent of the surveyed teachers perceived NCLB unfavorably.⁹ Although teachers described some benefits from NCLB policies in the classroom, such as an increase in teacher expectations of student learning, they also expressed concern about the costs, such as a narrowed curriculum. The authors concluded that of the teachers who had an unfavorable opinion of NCLB, chief among their concerns was the policy’s “lack of understanding and respect for the skills and experiences of teachers” (p. 6). In other words, these teachers perceived NCLB as a threat to their professional status and, as an extension, a threat to their job satisfaction, discussed in the next section.

Teacher autonomy and job satisfaction. Research indicates that teachers derive satisfaction in their work from many sources, including student achievement, professional

⁹ Guggino and Brint (2010) cautioned readers that their sample was not representative of NBCTs in California because of sampling design limitations and resulting oversampling of teachers from schools with higher average performance and lower percentages of students enrolled in free lunch programs.

development, and autonomy in the classroom. Using varying methods of investigation, the studies cited in this subsection focus on the intersection of teacher autonomy as it relates to job satisfaction. Consistently, scholars find that autonomy defined as control in the classroom can lead to increased job satisfaction and commitment to the workplace. For example, in a case study based on comparisons of teacher-reported perceptions of autonomy and job satisfaction across 12 high schools in six urban school districts in California, Florida, and New York, Archbald and Porter (1994) found that on average, teachers reported a higher perception of control in more decentralized environments and reported a high perception of pedagogy control regardless of whether their environment (i.e., departments, schools, districts, and states) was centralized or decentralized. The authors' overarching conclusion was one of caution. They contended that although in their study, teachers held a great deal of autonomy over classroom practice, future policies (e.g., NCLB) that might seek to "shape day-to-day content decisions" might indeed reduce teacher autonomy and, in turn, teacher job satisfaction.

Archbald and Porter (1994) used the "shrinking autonomy" construct to frame their study on the relationship between teacher autonomy and job satisfaction. According to the authors, proponents of local control assume that teachers and local authorities would readily and consistently strive for improved student achievement through better decisions than those presented by state and federal actors. In contrast, proponents of central control view the disparate nature of local control as anathema to academic standards and accountability. Characterized as having high, medium, or low curriculum control, teachers from each school reported on how their scope of influence over the following elements affected instruction: state and district curriculum guides, state and

district tests, textbooks, instruction, and departments. The study found that effectively, teachers vary in their perception of control over *what* content is taught but share a strong perception of control over *how* content is taught.

A second 1994 study found that “satisfied teachers report having more professional autonomy and challenge” (p. 10) and concluded that autonomy can be an important inducement for teacher recruitment and retention. Specifically, Kim and Loadman (1994) performed secondary data analysis on survey data from over 2,000 teachers to identify predictors of teacher job satisfaction. The authors used step-wise multiple regression and found that seven job satisfaction-related variables explained nearly 75 percent of the variance in perceived teacher job satisfaction. They grouped the indicators into extrinsic and intrinsic reward categories and found that an important intrinsic reward was autonomy in the classroom. Similarly, Brunetti (2001) surveyed 169 California high school teachers and documented the importance of teacher autonomy and its role in job satisfaction and, based on his research, speculated that a “deleterious effect on [teachers’] perception of professional autonomy” (p. 65) would result from standardized testing and curricula, thereby decreasing teacher job satisfaction. A quantitative study of nationally representative SASS data from the 1990–91 academic year underscored the importance of professional autonomy. The study found that among public school teachers, higher levels of perceived autonomy and influence were associated with high levels of commitment to teaching (Ingersoll & Alsalam, 1997). This finding held at the school level as well. That is, schools with higher percentages of perceived faculty autonomy had higher levels of faculty commitment to the school.

More recently, a case study of 39 elementary teachers found that 46 percent of teachers felt that NCLB had negatively influenced their morale specifically, with 80 percent of the teachers commenting that they had considered leaving the profession based on changes brought about by the Act (Deniston & Gerrity, 2010). However, the authors noted significant differences in the findings based on teachers' years of experience. Specifically, the teachers from the first elementary school who had the lower overall impression of NCLB had more years of experience in teaching than those in the second elementary school. Further, Deniston and Gerrity (2010) found that as years of teaching experience increase, a teacher's overall impression of NCLB decreased. In sum, the above studies highlight the potential for SBA reform, specifically NCLB, to affect teachers differently based on school context and teacher characteristics.

Teacher autonomy and retention. Extensive scholarship examines how teacher labor markets function, specifically with regard to teacher attrition and mobility (see for example, Margolis, J. (2008) and Guarino, C., Santibanez, L. and Daley, G. (2006)). Scholars identify an array of inputs that teachers consider as they make retention decisions. Principal among these inputs are economic incentives, professional development opportunities, workplace conditions, and societal perceptions of teaching (Hanushek, Kain, & Rivkin, 2004; Ingersoll, 2004; Johnson, 2006). Empirical studies that include autonomy as a condition of the workplace offer insight into the specific role that teacher autonomy plays in teacher retention. In their review of retention literature on "in-service policies," Guarino et al. (2006) cited that in addition to administrative support, mentoring programs, and class size, "the level of autonomy granted to teachers. . .often

appeared to play a prominent role in teachers' decisions to quit or remain on the job" (p. 183).

Findings discussed in this section permit three key generalizations. First, it seems that teachers who seek autonomy in their work are more likely to leave the profession if they do not find it. Second, teachers perceive autonomy and its relationship to other workplace conditions differently based on their own characteristics and those of the organization in which they work. Third, SASS informed much of the recent empirical research on teacher autonomy and retention. For example, using SASS data, Weiss (1999) concluded that first-year teachers "who experience autonomy and discretion are those who feel they have a say in developing curriculum, in selecting texts, content, teaching techniques, or discipline methods" (p. 865). More recently, and also using SASS, Ingersoll (2003) found that nearly half of the 50 percent of teachers who left the profession in the first six years cited exclusion from decision-making as a point of dissatisfaction. Ingersoll (2003) reported that the variables associated most strongly with commitment were perceived school leadership and teacher autonomy.

Scholars found similar findings from other data sources as well (Allensworth et al., 2009). Using data from teacher personnel records from 2002–03 to 2006–07, and data from principal, teacher, and student surveys, the Consortium on Chicago School Research (CCSR) analyzed teacher retention in Chicago Public Schools (CPS) during the spring semesters of 2001, 2003, and 2005. The authors state that, similar to the national rate of 84 percent, about 80 percent of CPS teachers return in the fall to the school where they had worked the previous spring. However, the authors point out that single-year rates do not reflect the true picture of CPS attrition. Over the 5-year period from 2002 to 2007, the

typical CPS school may lose more than 50 percent of its teaching force. Allensworth et al. (2009), the CCSR researchers, found that novice teachers left their teaching assignments at higher rates than did veteran teachers, and that this exit of novices from elementary schools was more pronounced than the exit from secondary schools. Specifically, in 2006-07, 67 percent of novice teachers stayed in their schools compared to 82 percent of veteran teachers, but at secondary schools about 75 percent of novice teachers compared to 80 percent of veteran teachers returned. Over the five-year period from 2002 to 2007, about 33 percent of novice teachers stayed in their schools.

In a hierarchical logistic regression of teacher survey data, the CCSR team established a link between “teacher influence” and retention. In the study, the measure of teacher influence over the selection of instructional materials and techniques overlaps with the measure of teacher-perceived autonomy used in this dissertation. The teacher influence measure varies because the measure also includes other indicators of influence over in-service program planning and teacher hiring. Regardless, the study found that “most important for teacher stability is the degree to which teachers feel they have influence over school decisions. In both elementary and high schools, stability rates were at least five percentage points higher in schools with substantial teacher influence, compared to schools where teachers had little influence over their work environment” (p. 26). Further, the study identified a strong relationship between teacher retention and school leadership such that teachers were more likely to stay in schools with “inclusive leadership,” defined as teachers having influence over working environment and having trust in their principal’s capabilities to lead the organization. In sum, the CPS study indicates that novice teachers may behave differently than do veterans; that teacher

influence over their work has a role in teacher retention, and that an association exists between teacher impressions of their principals and both autonomy and retention.

In contrast to the findings for novice teachers from the CPS study—an earlier study that investigated the link between school reform, teacher retention, and teacher perception of autonomy—found that perceived autonomy was more closely linked to work-related variables than to academic and training experiences, years as a teacher, or age and gender (Pearson, 1995). Pearson (1995) used blockwise (or stepwise) multiple regression to predict teacher autonomy from three groups of predictor variables: work-related variables (e.g., instructional load, job stress, satisfaction with salary), attitudinal variables (e.g., perception of parental involvement, perception of administrative support, and perception of students), and teaching level (elementary and secondary). The scholar found that job satisfaction, perceived paperwork load, positive attitude towards students, and teaching level are associated with autonomy. Similar to the CPS study, Pearson (1995) found differences based on grade level taught. Specifically, teachers in secondary schools reported a higher perception of autonomy than teachers in lower grades.

More recently, Boyd, Grossman, Ing, Lankford, Loeb, and Wyckoff (2011) studied teacher retention in New York City schools and built a statistical model that included a factor of teacher influence over their work as an independent variable in the model to predict teacher retention. The key finding from this study concerned administrative support as a significant and critical element of teacher mobility decisions. Although this study did not find teacher autonomy to be the biggest driver of retention decisions, the authors recognized a need for studies that do investigate teacher autonomy, and they recognized that part of administrative support might be the extent to which

school leaders foster or inhibit teacher autonomy in the classroom. Consequently, this research offers two critical contributions to the proposed study. First, the authors modeled their factor of teacher influence using five of the six questions from the SASS teacher questionnaire that I propose to use in my study and set a precedent for these measures as a proxy for teacher autonomy in the classroom. Second, the authors cite the advantage of SASS data as nationally representative and as a useful data source for teacher research.¹⁰

Ingersoll and May (2010; 2011) published two recent studies that examined the specific relationship between teacher autonomy and retention. In 2010, the research team investigated the teacher- and school-level factors that led to the turnover of mathematics and science teachers. Ingersoll and May (2010) used five waves of SASS and TFS data to create a descriptive analysis and used 2003–04 SASS and 2004–05 Teacher Follow-up Survey (TFS) data to create their multilevel logistic regression models. The dependent variable was a dichotomous variable that indicated whether the teacher had remained in the same school as in the prior year. The authors used factor analysis to create a “Classroom Teacher Autonomy” variable that they included as part of the organizational characteristics of schools. Specifically, Ingersoll and May (2010) used SASS questions about teacher perception of control in their classrooms across six areas to build the teacher autonomy measure: textbook and classroom materials selection; content, topics, and skills to be taught; teaching technique selection; grade assignment and evaluation of students; homework assignment; and student discipline.

¹⁰ The authors state that their study is not suited to SASS data because of the potential for common-source bias that may occur in a study on teacher retention that uses both SASS and Teacher Follow-up Survey (TFS) data because the same teachers respond to SASS and TFS. This common-source bias is not a concern for the proposed study because I do not propose to use TFS data.

The study found mixed results for autonomy and turnover. The authors stated that “by far the strongest predictor for mathematics teachers was the degree of individual classroom autonomy held by teachers; in contrast for science teachers classroom autonomy was a non-factor” (p. 44). It is worth noting that this study occurred prior to the science proficiency requirements of NCLB that begin in 2014. Nonetheless, this contrasting finding by subject indicates that not only do different teachers perceive autonomy differently, but also that subject matter taught may be an important factor to consider when examining teacher perception of autonomy.

In a similar study, Ingersoll and May (2011) sought to explain recruitment, employment and retention patterns of minority teachers. Again, the authors used data from multiple waves of SASS data to describe the trend of minority teacher turnover and used the 2003–04 SASS and 2004–05 TFS wave of data to identify predictors of minority teacher turnover. The authors created the teacher autonomy constructs in the same manner as the 2010 study. In this study, the authors described a more uniform finding in that “schools with higher average levels of individual teacher’s classroom autonomy had lower levels of turnover” (p. 35).

The Ingersoll and May (2011, 2010) studies inform this dissertation in several ways. First, the authors emphasize the utility of SASS as valuable data from which to glean nationally representative findings about teachers and their working conditions—of which autonomy is a key element. Second, the authors performed their analyses using statistical procedures that account for the nested, or multilevel, structure of the data, as I intend to do. Third, the authors used responses to the six measures of teacher classroom control that this study also uses because these questions most directly tap into the notion

of teacher autonomy in the classroom. Fourth, the authors demonstrate that retention and autonomy relate in a statistically significant manner, and conclude that autonomy is a construct that merits further exploration because of this empirically established relationship.

Teacher autonomy and SBA reform. As detailed earlier, the 1980s marked a different era of control over the classroom as a result of SBA reform (DeBray, 2006; Eisner, 1995; Jennings & Rentner, 2006; McDonnell, 2004; Sunderman & Kim, 2007; Superfine, 2005). As early as 1981, scholars identified a tension between teacher autonomy and accountability policies enacted at the school level. For example, Leiter (1981) examined the relationship between teacher autonomy and accountability through a lens of “competition on claims of school goals” (p. 236) to identify the variation in teacher autonomy across three school settings—rural, suburban, and urban. In his case study, Leiter (1981) found that the competition of claims over the suburban school was lower than it was in rural and city schools. In the suburban schools the principals held influence over teachers in terms of attendance, punctuality, and personal behavior (i.e., behavior outside of the classroom that reflected moral values consistent with those of the principal), but did not interfere with day-to-day classroom instruction. Conversely, in the rural schools, the principals had an “upward flow” of information on teacher practices in the classroom. This unidirectional flow of information created uncertainty among teachers who did not know what information the principal had or how the information would be used and diminished teacher sense of efficacy and, in turn, autonomy in the classroom.

In the city schools the principals focused more on managing external demands of federal and state accountability related to equity and accountability programs and did not interfere with instruction. Leiter (1981) found that in the city schools, teacher perceptions of their own autonomy reflected their perception that the principal lacked autonomy and efficacy. The “control vacuum” developed because to teachers the principal appeared to

be powerless over the school and ineffective in his leadership role. In turn, teachers perceived themselves to be unable to achieve change or progress in the classroom, which resulted in low perceptions of teacher autonomy. Leiter (1981) concluded that competitive claims on school goals diminish principal and teacher autonomy in schools and classrooms because when teachers perceive principals as having low autonomy and efficacy, they appear more likely to perceive low autonomy for themselves.

More recently, other researchers have documented the pattern of decreased teacher autonomy in the classroom associated with increased claims on school goals. Russell and McCombs (2006) conducted a study of over 3,000 teachers in California, Georgia, and Pennsylvania. The authors investigated school administrator and teacher perceptions of standards, assessments, and state accountability systems. The study found that teachers in California and Pennsylvania who taught in Title I schools¹¹ were more apt than those in non-Title I schools to focus on student achievement because of the incentives and sanctions of AYP status.¹² They also found that over 40 percent of teachers in California and Georgia, and over 60 percent of teachers in Pennsylvania reported a negative effect from the reforms because they needed to “teach to the test” and because teachers would carry the blame if the school failed to meet adequate yearly progress (AYP).

In the early 1980s, very large urban districts and some states became more directly involved in curriculum in efforts to achieve some level of alignment between

¹¹ Title I schools are those that receive federal funding based on the percentage of students in the school considered to be from impoverished families. These schools received federal funds based on an allocation formula and are subject to all NCLB requirements and sanctions.

¹² Teachers in Georgia also focused on student achievement because state policy calls for sanctions and rewards based on AYP status for non-Title I schools as well as Title I schools. Thus, findings show that, AYP status impacted teacher motivation regardless of Title I status.

standards and curricula. Rowan (1990) describes educator reaction to curriculum alignment as a burgeoning shift in curricular control that “undermined professional autonomy of teachers (p. 353). According to Rowan, the theory of action behind standards was that “education ‘technology’ was amenable to routinization” (p. 356) and that districts and state-level actors could script for teachers. A 2007 study of novice teachers¹³ who work in New York City Public Schools confirms Rowan’s statements. Crocco and Costigan (2007) spent more than five years interviewing novice English and social studies teachers. The authors reported that across the sample, teachers described “shrinking space for their classroom-based decisions making” (p. 521) due to scripted lessons. The effect was particularly acute for middle-school teachers and those with the strongest credentials.

The authors are careful to point out that novice teachers sought autonomy in concert with support from mentors and school administrators. Nonetheless, Crocco and Costigan (2007) concluded that novice teachers perceive NCLB as reducing their ability to exercise control over the classroom and suggest that

the problems [novice teachers] prescribe are pervasive, and the themes articulated around issue of autonomy, lack of proper support, and frustration with the narrowing of curricular and pedagogical options so widely shared that we do not believe them to be idiosyncratic symptoms of underprepared or immature new teachers. Instead we believe that the problems reflect a system that has been established to make schooling ‘teacher proof’ in a misguided effort to increase student achievement on high-stakes tests. (p. 528)

Put another way, the authors contend that novice teachers may be the voices in this case study of the relationship between SBA reform and teacher autonomy, but

¹³ In this study the authors define novice teachers as those who have at least several months but no more than five years of experience.

given the systemic approach to SBA reform, it is likely that more-experienced teachers would echo the novice teacher sentiments.

On another front, early scholars of SBA reform sought to specifically define and quantify teacher autonomy, and ultimately developed the Teacher Autonomy Scale (TAS) to measure general teacher autonomy and curricular teacher autonomy. Effectively, through TAS two studies from 1993 began the quantitative study of the relationship between teacher autonomy and education reform policies. The main contribution of Pearson and Hall (1993) was their rigorous approach to the development of an autonomy construct through factor analysis. These authors were able to establish autonomy as a latent construct based on observable items. Specifically, Pearson and Hall (1993) conducted two studies to build a measurable construct for teacher autonomy called the Teacher Autonomy Scale (TAS). The authors cited “restructuring school” reforms as a mechanism that elevated teacher autonomy on the policy landscape and recognized that quantitative measures for teacher-perceived autonomy were in their nascent stages in the early 1990s.

Pearson and Hall (1993) sought to “validate an appropriate instrument for measuring teacher autonomy” to measure “perceptions that teachers have regarding whether they can control their work environment” (p. 173). The goal of the first study was to refine a 35-item scale and develop a “psychometrically sound” (p.173) instrument that could be administered in the second study as a basis for examining relationships between teacher perception of autonomy and various demographic characteristics of teachers. The result from the first study was a 20-item instrument intended to measure four aspects of teacher autonomy:

- Selection of activities and materials
- Classroom standards of conduct
- Instructional planning and sequencing
- Personal on-the-job decision-making.

The new Teacher Autonomy Scale (TAS) had both negative and positive items and was internally consistent with strong item-total correlations.¹⁴ In other words, the scale captured the construct of autonomy.

The goal of the second study was to administer the TAS to a sample of public elementary, middle, and high school teachers and draw some empirically based conclusions about teacher autonomy. The authors determined that teacher autonomy was not a signal construct, but had two distinct dimensions—general teaching autonomy and curricular autonomy. Perhaps contrary to what one might expect, the findings of this study suggested that teachers did not vary on their perception of autonomy by gender or education level. Teachers did differ by grade-level taught. Those in middle schools perceived higher levels of autonomy than did those in elementary or secondary schools. This early study offered a starting point from which to advance the study of teacher autonomy through a quantifiable autonomy construct.

Since the 1990s and because of the progression of SBA reform, teacher perceptions of autonomy in the workplace have gained prominence as an area of study. For SBA reforms generally and NCLB specifically, scholars have found that teachers' scope of influence has narrowed while their scope of responsibility has widened (Burbules & Densmore, 1991a; Grant, 1983; Ingersoll, 2007). For example, a 2002 exploratory study of four science teachers in urban schools in Alabama and Ohio offers anecdotal and suggestive findings regarding assessment and accountability. The four

¹⁴ Internal consistency coefficient of 0.91 and no item-total correlation less than 0.44.

teachers stated their frustration with the inability to exercise the professional judgment they had acquired through their teacher education programs. They commented that the state test dictated the content of their lessons, in part because of the focus on the assessment and in part because of faulty alignment between state exams and on-site curriculum.

Like other SBA reforms, NCLB's implementation process uses external controls (e.g., state-mandated curricula and assessments) to hold teachers accountable for educational reform and increased student achievement (Finnigan & Gross, 2007).

Teachers face an ever-increasing set of demands in the classroom, from HQT requirements to adoption of "scientifically based" curriculum to prescribed instructional methods, (McDonnell, 2004; O'Day, 2002; Sunderman & Kim, 2007). Thus, as NCLB implementation stands today, it is not surprising that research on teacher perception of control and influence over the classroom reveals a tension between policy goals and teacher practices in the classroom. For example, Pearson and Moomaw (2005; 2006) conducted two studies that looked at autonomy as dependent on measures such as workplace conditions and teaching level. Specific to autonomy, the authors found that teacher perceptions of autonomy related to working conditions more than to pre-service training, demographic characteristics, or years of experience.

Researchers have also found that NCLB has centralized decision-making about daily activities in the classroom, specifically the content and the pedagogy (Santoro, 2011; Sunderman et al., 2004; Valli, Croninger, Chambliss, Graeber, & Buese, 2008). For example, over a four-year period, Valli and Buese (2007) examined elementary teachers who operated under NCLB school-level accountability policies to investigate what

changes, if any, in classroom practice may be attributed to NCLB. The authors used differentiated instruction as an indicator for change in the classroom. After examining math and reading instruction, the authors concluded that teachers' tasks increased in number and scope, but their autonomy decreased. Teachers had become "more hierarchically controlled" (p. 551).

The authors also offered an example of change regarding teacher collaboration. During the first year of the study, researchers observed that teachers collaborated with peers on how best to differentiate their students for improved learning outcomes. By 2004–05 researchers noted the teachers focused collaboration on how to differentiate instruction for testing strategies more so than to differentiate for specific education outcomes or learning goals. Further, "teachers were too busy learning how to read data to regroup and re-teach students. . .to engage in 'rich deliberations' about the substance of their teaching" (p. 552). The authors' observations suggest a shift in teacher perception of autonomy both for instructional practice and for peer-group interactions because in some instances teachers changed practice not based on teacher beliefs or professional expertise but on testing policies.

Mathison and Freeman (2003) conducted an ethnographic study of elementary school teachers in two New York City schools during the 2001–02 school year to explore teacher responses to the New York State testing regime. The authors described teachers as perpetually conflicted actors caught between the desire to administer what they perceived as appropriate instruction and what they perceived as teaching to the test. The authors stated,

Teachers must often do the wrong thing in order to do the right thing, sort of. The means are approaches to teaching and content that teachers might not chose—that do not represent good professional practice and, the state's desired ends

(high test scores) are a poor but powerful proxy for the teachers' desired ends (the contextually appropriate success of every child). (p. 18)

Further, the authors declared that these two sets of teachers neither sought “complete autonomy” nor disagreed with centralization of some aspects of teaching. However, teachers did have the impression that they needed to turn away from their training to achieve the stated goals of state-mandated and school-implemented testing. Although the authors do not draw causal conclusions from this ethnography, it appears that the testing policies negatively influenced teacher perception of autonomy in the classroom.

Specific to the AYP element of NCLB, a recent study of teachers in South Carolina found that working conditions, including teacher empowerment (indicated by their level of participation in decision-making regarding matters that affected students), predicted AYP status (Hirsch, 2005).¹⁵ The authors found that the more empowered the teachers, the greater the likelihood of meeting AYP goals. Though not specific to autonomy, recent studies have begun to uncover how AYP status can affect teacher perceptions of the workplace. For example, Sunderman, Tracy, Kim, and Orfield (2004) conducted a study that divided responses along AYP status lines—schools that met adequate yearly progress and schools that were in need of improvement for one, two, or three years. The researchers surveyed nearly 1,500 teachers from two school districts across two states (80 schools in Fresno, California and 49 schools in Richmond, Virginia). All schools had at least three grades that required testing by NCLB. To capture perspectives across levels of AYP status, the sample included teachers from schools that were in their first, second, or third year of improvement status, or that had met adequate yearly progress.

¹⁵ The study reported responses from 15,200 teachers across 90 percent of schools and 100 percent of districts.

In the study, student characteristics of the schools matched the overall district population with two important exceptions along racial and economic lines. First, compared to the district wide average, a higher percentage of students in the schools in need of improvement in Richmond were black (89.5 percent vs. 95.7 percent, respectively) and a lower percentage were white (7.4 percent vs. 2.1 percent, respectively). Second, a higher percentage of students in the schools in need of improvement were low-poverty (95 percent vs. 81.2 percent, respectively). The authors found that teachers in both types of schools (i.e., in need of improvement and adequate progress met) had a similar distribution of views across the 5-point Likert-scaled questions. However, two areas of difference were teachers' perceptions of themselves and of their students. When asked their perceptions regarding teacher commitment to improving student achievement and to providing high quality instruction, more teachers in schools that met AYP reported they "strongly agreed" with the statements. The differences were between 10 to 20 percentages points. While not nationally generalizable, this study does support the notion that teachers differ across school contexts (i.e., AYP status) in their perceptions of and, perhaps their experiences with, autonomy in the classroom.

Similarly, Olson (2008) conducted a case study of six English teachers in their first year of teaching during the 2005–06 academic year to investigate novice teacher identity. In addition to finding that the six teachers' identities changed depending on their students, their school environment, and their personal expectations, Olson (2008) found that

the administrative tendency to give new teachers academically low-tracked courses, and these teachers' own social justice desires to find communities where they believed they could most 'make a difference' meant they were not teaching

English the way they expected. . . . Their apprenticeship of observation about what it meant to be a teacher (occurring before *No Child Left Behind*) very likely did not include top-down policy mandates, prescriptive curricula, and strict teacher accountability measures. They had expected to have autonomy and control in their classrooms and had been professionally prepared to teach as they thought best. (p. 33)

In contrast, Goldstein's (2008) case study of four kindergarten teachers from two suburban high-performing elementary schools in Texas documented the transformation that kindergarten teachers experienced regarding SBA reform and autonomy in the classroom. Perhaps unexpectedly, the kindergarten teachers were able to maintain some level of decision-making about how to teach the state-mandated curriculum. All four teachers (with teaching experience ranging from five to 28 years) stated they were accustomed to standards-based "curricular constraints" (p. 458). The author found her participants were able to maintain some degree of street-level policymaking because they were able to shape state-mandated curriculum such that "every policy decision was unique and deliberate and reflected attention to obligations, desire for autonomy, and the issue of professional discretion" (p. 473).

Goldstein (2008) noted that the teachers' ability to maintain their perceived autonomy may not be the case in low-performing schools and suggested that, based on evidence in this case study, teachers in low-performing schools would have less "decision-making autonomy" than the four teachers in the case study. Again, research indicates that teachers in a standards-based classroom perceive autonomy differently based on their levels of experience, perceptions of their leader, school context, and student characteristics. The gap that remains in the research is how much AYP status might directly explain these differences and how much the most recent wave of SASS data might inform the variations in teacher and school characteristics.

Implications of Literature on Teacher Autonomy for the Present Study

Autonomy in general and teacher autonomy in particular are complex and much-debated constructs (Burbules & Densmore, 1991a; Campbell, E. 2006; Elmore, R. F., 2002; Ingersoll, 2007; L. M. McNeil, 1988). Scholars understand that autonomy runs on a “continuum spanning from complete constraint to total freedom” (Gawlik, 2007, p. 526). The absence of autonomy may strangle teacher motivation and creativity, but complete autonomy could make for inconsistent learning opportunities for students and unruly governance of the education system. Further, components of the public school system, such as professional associations and bureaucratic structures, serve as systemic checks on the autonomy of public-school teachers (Glazer, 2008; Leiter, 1981; Retsinas, 1982). Notwithstanding the debate over how much teacher autonomy is appropriate in the classroom, the research documented above establishes that teachers value and expect autonomy in the classroom. Scholarship further establishes that teachers perceive the effects of SBA reform on their autonomy in the classroom differently based on teacher-level and school-level characteristics, such as

- teachers who teach in high-poverty schools versus lower poverty (Diamond & Spillane, 2004; Goldstein, 2008; Santoro, 2011);
- teachers who instruct math or science versus other subjects (Ingersoll & May, 2010; Settlage & Meadows, 2002);
- teachers who instruct elementary versus middle versus secondary schools (Pearson & Hall, 1993); and
- teachers who are new to the profession versus those who are veterans (Crocco & Costigan, 2007).

The studies cited in this literature review offer six key takeaways that inform the present proposal. First, teacher-perceived autonomy varies based on teacher characteristics and school context. Second, some studies used the construct of autonomy as a dependent or outcome variable, and others used it as predictor variable. Regardless, where applicable, the studies indicated that autonomy is important to teachers. Third, reminiscent of Dahl's (1984) power-influence model, teachers' perception of autonomy varies based the domain under scrutiny, as well as the comparative power held by other players on the field. That is, teachers do not perceive their control in a vacuum – or behind a closed door. They operate within a larger set of power structures that inform their perceptions of autonomy. Fourth, SASS is an established and relied-upon source of data to use when exploring questions of teacher autonomy. Fifth, the distinction between different levels of data, such as school and teacher, is an essential methodological step to capture teacher-perceived autonomy because scholars have established that different levels within the school system influence teacher autonomy differently. Finally, the findings from the studies reviewed here are helpful because they establish a baseline for teacher-perceived autonomy and expose the factors that influence those perceptions, such as school leadership, student demographics, and AYP status. Informed by the prior research reviewed in this chapter, the next chapter details the two empirical frameworks and multiple statistical procedures I use in the current study.

Chapter 3: Methods

This chapter has five sections. The first section discusses the details of the data and the data collection process for the Schools and Staffing Survey (SASS). The second section presents the two empirical frameworks of the study. The third and fourth sections detail the analytic samples and variables, respectively. The fifth section presents the statistical procedures of the study.

Description of Data Source

Since the 1980s, the National Center for Education Statistics (NCES) has collected data on elementary and secondary school systems through the Schools and Staffing Survey (SASS). Originally, the surveys focused on “teacher demand and shortage, teacher and administrator characteristics, school programs, and general conditions in schools” (<http://nces.ed.gov/surveys/sass>). As the survey program has matured, NCES has added other areas, including teacher perceptions of the workplace and overall school performance. In broad terms, NCES intends SASS to offer researchers robust data to use to investigate an array of issues concerning U.S. elementary and secondary schooling. As of 2007–08, SASS consisted of five surveys: School district, School, Principal, Teacher, Library/Media center.¹⁶

One advantage of SASS is that it allows multiple collection points over time. NCES states that since the surveys repeat questions across administrations, researchers can use the data to examine trends over time for policy areas of interest

¹⁶ In 1993–94 the District file was the “Teacher Demand and Shortage” file and a sixth file “Student Records” was administered. Beginning in 1999–2000, NCES changed the “Teacher Demand and Shortage” to the “School District” file and discontinued the student data file altogether.

(<http://nces.ed.gov/surveys/sass/issues.asp>). To that end, the proposed dissertation examines data from four administrations of SASS across three surveys:

- SASS 1993–1994, 1999–2000, 2003–04 : Teacher and school files
- SASS 2007–2008: Teacher, school, and principal files

Descriptions of all data sources come primarily from their respective documentation (ED, 1996, 2004, 2007, 2010). SASS uses schools as the primary sampling unit. The Common Core of Data (CCD) is the sampling frame from which NCES choose public schools for participation. The sampling design for SASS is a stratified probability proportionate to size sample. That is, initially NCES identifies a sampling frame of schools based on school type (e.g., charter school, high Alaskan Native population). To ensure support of statistical outcomes based on characteristics of interest, the sample draws from some schools at a higher rate than others if their particular characteristic is uncommon across the population. For example, SASS includes all schools funded by the Bureau of Indian Education (BIE) in the sampling frame because there are so few BIE schools relative to the universe of publicly funded schools. Once the school sampling frame is complete, NCES identifies a sample of teachers and all principals and all school districts associated with the school sampling frame for inclusion in the SASS data collection.¹⁷

NCES has made changes to SASS with each administration; however, similar structure and data collection techniques undergird the four administrations discussed here. The most obvious change is the sample size, which has decreased since the 1993 administration. Nonetheless, each data set supports state and national estimates for public elementary and secondary schools, teachers, and principals. Table 2 displays the sample

¹⁷ Some years have more schools than principals. This is due to principal non-response to the principal survey.

sizes for schools, teachers, and principals¹⁸ for each administration and the associated unweighted response rates.¹⁹ NCES conducted non-response bias analyses for all SASS data. For years 1993–94,

Table 2

Sample Size

	1993-94		1999-2000		2003-04		2007-08	
	Sample size (n)	Response rate (%)	Sample size (n)	Response rate (%)	Sample size (n)	Response rate (%)	Sample size (n)	Response rate (%)
Schools	8,770	92.0	8,430	88.5	7,990	81.9	7,570	80.5
Teacher	47,100	88.9	42,090	81.2	43,240	84.0	38,240	84.0
Principal	n/a		n/a		n/a		7,460	79.5

1999–2000, and 2003–04, NCES indicated that the non-response rates did not pose a substantial bias in SASS estimates. In 2007–08, NCES indicated that “potential bias remains” for secondary schools in Nebraska and for the “central city” schools, but that the final 2007-08 weights reduce this potential bias.

SASS data files are fully imputed datasets built from the responses to SASS questionnaires.²⁰ Data collection runs from October through June of the specified school year and incorporates both pen and paper questionnaires as well as telephone and in-person follow-up interviews. Descriptions of the surveys relevant to the dissertation proposal are below.

¹⁸ The proposed dissertation will use the principal file from 2007–08, thus principal counts for other years are not shown.

¹⁹ To comply with NCES standards for restricted-use data, I have rounded all sample sizes to the nearest 10.

²⁰ According to NCES, the imputation process for all SASS datasets involves a two-step process. In the first stage, NCES staff interpolates responses based on actual responses to related items in the same questionnaire or related items from other questionnaires that were completed by the same respondent. The second stage of imputation may involve imputing data based on the mean or mode calculated from responses of similar respondents. Subsequent to the imputation procedure(s), a Census Bureau analyst reviews the imputed data to correct any resulting outliers or inconsistencies in the data.

School file. NCES collects school data through the SASS school questionnaire. This survey is unique in that several school administrators or staff may respond to parts of the survey depending on the question. The principal, or designee, may answer the majority of questions but seek responses from other administrative personnel (e.g., school counselor) on such topics as enrollment, school programs, staffing, and Title I status.

Principal file. Concurrent with the school data collection, SASS collects principal data through the principal questionnaire. NCES contacts every principal contained in the school sample. This survey collects information specific to the school leader, such as training and experience, and information regarding the school as a whole, such as academic performance.

Teacher file. NCES draws a sample of teachers from the schools in the sampling frame. The teacher survey contains questions about teaching position (e.g., full-time, part-time, grades and subject matter taught), number of students taught, working conditions, attitudes and perceptions of the workplace, and various demographic characteristics (e.g., educational background, employment background).

Empirical Frameworks

The hypotheses that drive my empirical models are that teacher perceptions of autonomy have changed over time and that SBA reforms are associated with that change. Specifically, I hypothesize that for 2007–08 particular teacher and school characteristics affected teacher perceptions of autonomy, such as grade level and subject matter taught, perception of the school leader, years of experience, school level, and Title I status. More specifically, I posit that AYP status, as the direct measure of accountability, negatively associates with teacher perception of autonomy. I theorize that teachers who work in

schools that do not meet AYP have lower or more varied perceptions of autonomy compared to their peers who work in schools that made AYP. I further posit that teacher perceptions vary in schools that face sanctions because of their Title I status, or in schools that operate in states that apply sanctions to any school that has failed AYP regardless of Title I status.

Prior to the investigation of specific teacher and school characteristics in 2007–08, I investigate any change in the degree of teacher-perceived autonomy in the classroom over time. Figure 2 illustrates the relationships I explore using each of the four SASS administrations. The domain and scope of teacher-perceived autonomy remain constant over time. Again, the domain is the classroom and the scope consists of six SASS questions related to classroom control as perceived by teachers: textbook selection; teaching technique selection; content, topics, and skills selection; student grading and evaluation; student discipline; and homework determination. In contrast, the degree of autonomy fluctuates based on responses to those six questions as measured with the autonomy scale (i.e., the dependent variable that I describe in detail later in this chapter). The descriptive analyses produce a portrait of the degree of teacher-perceived classroom autonomy vis-a-vis the central tendency (mean), total variation (CV), and source of variation (ICC) in the degree of teacher-perceived autonomy in the classroom domain.

As described in the discussion of the conceptual framework, arrows *a* and *b* represent the relationship between the SBA reform policies that emanate from the federal, state, and district levels of the education system and the teacher and school levels. However, the grayed representation of these arrows illustrates that although these

relationships exist in the system, I am not able to capture their effects in the empirical analyses based on the data available for the current study. Arrow *c* represents the direct effect that school characteristics have on teacher-perceived autonomy (e.g., school performance, Title I status, school leader). Arrow *d* represents the indirect effect that schools exert over teacher-perceived autonomy through administrative actions that impinge on the teachers themselves, such as scheduling, room assignment, accountability programs, and any number of organizational factors that affect teachers' workplaces, as previously discussed in the literature review section of this dissertation. Arrow *e* represents the relationships between teacher characteristics and their perceived autonomy. The over-time descriptive models do not include any specific characteristics at either level because the focus is to identify any pattern or trend in autonomy values from 1993–04 to 2007–08. The descriptive analyses include calculation and comparison of means, coefficients of variation (CV), and intraclass correlation coefficients (ICC) for the autonomy scale without any mediating variables and serve to set the context for the 2007–08 HLM analyses. The descriptive statistics enable me to identify overall changes in autonomy, in terms of mean values of the autonomy scale, as well as total variation in the scale and changes in the sources of that variation—between school (level 2) and teacher (level 1) components—over time.

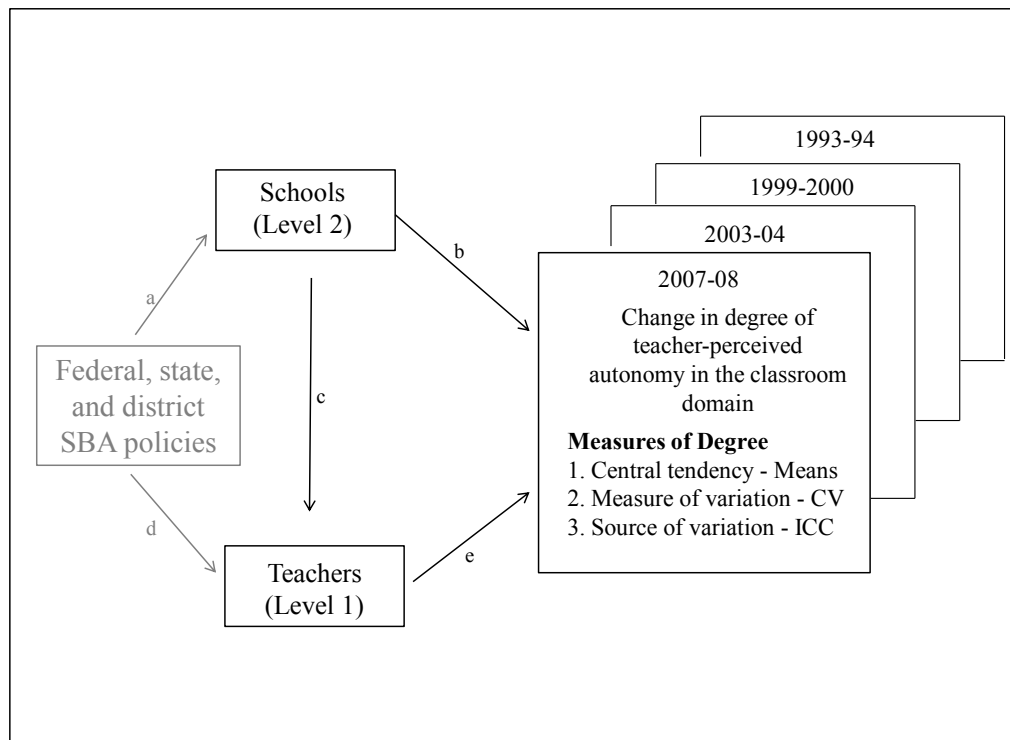


Figure 2. Empirical model of change in degree of teacher-perceived autonomy in the classroom domain: 1993—2007.

Figure 3 illustrates the relationships that I explore through multilevel regression analyses techniques on 2007–08 data (HLM). Unlike the second empirical model, the model in Figure 3 is specific to a point-in-time and includes specific indicator and control variables (i.e., school and teacher characteristics) to model the relationships. Consistent with the conceptual model, this model illustrates the relationship between teacher and school characteristics and the scope and degree of teacher-perceived autonomy in the classroom domain. I have added a relationship to this empirical model (arrow *f*) to illustrate that school-level characteristics may influence the relationship between teacher characteristics and teacher-perceived autonomy in the classroom. That is, the

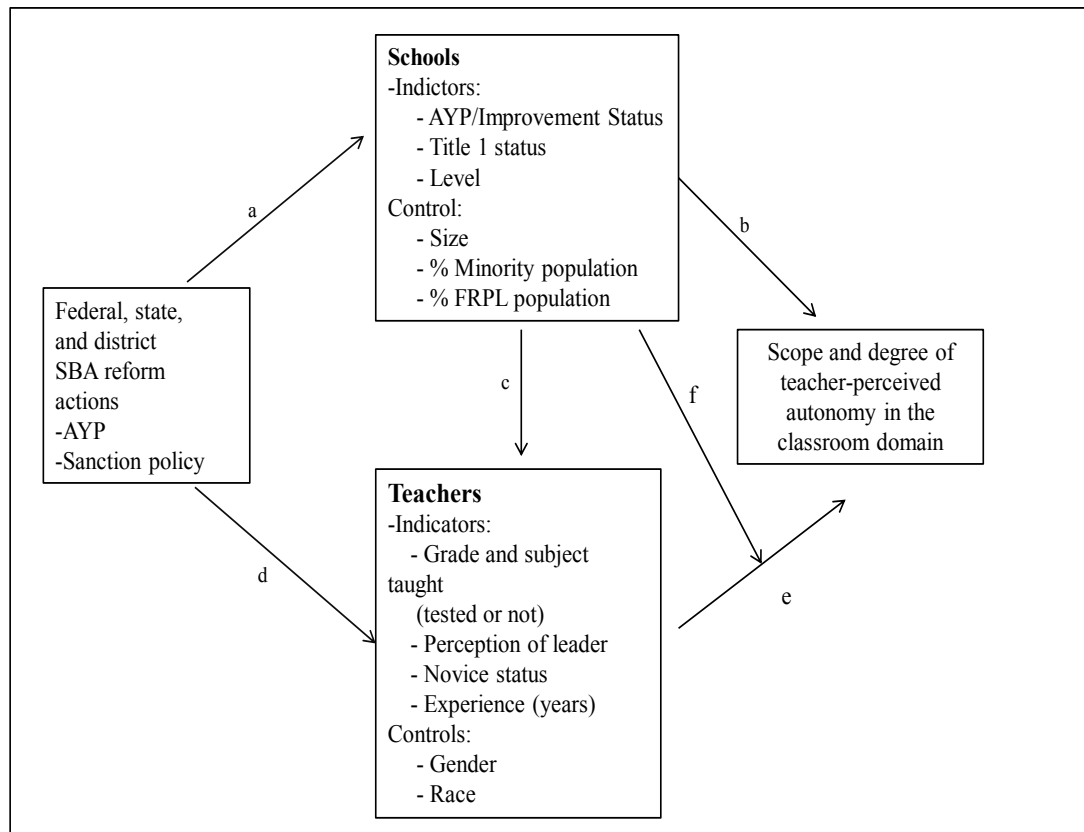


Figure 3. Empirical model of the relationships between SBA reform, school and teacher characteristics, and the degree of teacher-perceived autonomy in the classroom domain.

relationships between teacher characteristics and perceived autonomy may differ based on the type of schools in which teachers work. It is only through multilevel modeling that I am able to isolate the degree and the direction of influence that schools have on teacher-perceived autonomy.

Multilevel modeling allows me to measure how much schools influence teacher perceptions of autonomy over time (descriptively by examining change in the ICC values), and specifically how schools influence teacher-perceived autonomy using the full HLM model on 2007–08 SASS data. I model the relationships between school and

teacher indicators, with particular attention to AYP status as a measure of SBA reform at the school level and the state's policy on sanctions as a measure of specific SBA reform at the state level. Unlike the other characteristics in Figure 3, I list AYP as external to schools and teachers, as well as list AYP status as a school characteristic. I include AYP status in the federal, state, and district boxes and the school box because other levels of the public school system impose AYP status as an indicator of accountability, but schools implement action to achieve AYP. Thus, I consider AYP status as integral to multiple levels of the system. The remaining school and teacher characteristics are compositional in nature. Again, in model 3 the scope of autonomy is the standardized scale of autonomy based on teacher responses to six SASS items that elicited teacher perceptions of "complete control" to "no control" in the classroom.

My statistical analyses use the SASS data to investigate my research questions. My overriding interest is teacher-perceived autonomy in the classroom. However, teacher-perceived autonomy is a latent construct not directly captured in SASS. This construct presents some measurement challenges for the analyses that I detail below. A particular measurement challenge for these analyses is comparing average autonomy across years in a meaningful way. For example, initial descriptive analyses indicate that across all four SASS administrations teachers perceive high average levels of autonomy. Thus, my analyses will focus on the *variation* in the outcome variable over time. To illustrate, on a scale from 1 to 4 of autonomy (4 being high of autonomy), if teachers hover on average at 3.5, that would indicate a high level of autonomy and, more technically, a skewed outcome variable. Notwithstanding the 3.5 mean value, if over time the average remains high but more variation in the data emerges, that would indicate less

agreement among teachers on how they perceive autonomy. Simply put, more variation in the outcome measure equates to less agreement among teachers about perceptions of autonomy. This variation in perception may indicate a differential effect of SBA reform on teacher-perceived autonomy based on school context or teacher characteristics. For instance, teacher perceptions may vary for teachers in Title I schools more so than they do for those in non-Title I schools; they may vary based on the novice status of teachers who might have different expectations of their profession compared to veteran teachers who are unaccustomed to classroom practices based on NCLB accountability models of reform. In the paragraphs that follow, I describe the analytic samples, and the attributes of the dependent, indicator, and control variables in this study that enable me to investigate sources of variation.

Analytic Samples

Table 3 lists the analytic sample sizes for SASS by year and data file.²¹ The analytic samples do not contain the full complement of observations for several reasons. I limited the analytic sample in 1993–94, 1999–2000, and 2003–04 to schools with responding teachers. Thus, the samples do not contain schools without teacher-level data, nor do they contain teachers without corresponding school-level data. This limitation reduced the school samples by 3.2 percent 5.5 percent, and 2.2 percent, respectively and the teacher samples by 6.5 percent, 8.8 percent, and 6.8 percent, respectively. I did not perform missing data analysis on these samples because the final teacher weights in each of the three years account for bias that may result from school and teacher non response.

²¹ To comply with NCES standards for restricted-use data, I have rounded all sample sizes to the nearest 10.

Table 3

Unweighted Analytic Sample Sizes by Year

	1993-94	1999-00	2003-04	2007-08
School-level	8,480	7,970	7,820	6,580
Teacher-level	44,040	38,740	40,280	33,110

For 2007–08 analyses, I made three changes to the sample. First, I limited the school-level data to schools with at least 75 enrollees, which reduced the available school-level sample by 4.7 percent to 7,220 schools. Therefore, the results of my analyses do not represent schools with fewer than 75 students. Second, I limited the school-level sample to those schools with corresponding principal reports on AYP status, which eliminated an additional 4.9 percent to yield 6,860 schools. At this point I performed missing data analysis on school-level data between the base sample of 7,220 schools with at least 75 students, the level-2 sample of 6,860 after limiting for principal response, and the 360 excluded cases from the principal sample. Table 4 details the results of the analysis.

At a p-value of 0.05, independent t tests yielded three statistically significant differences between the level-2 and the principal samples. The level-2 sample has lower percentages of minority students and students who qualify for free- and reduced-price lunch, as well as lower average total enrollment. However, no differences are statistically significant between the base sample of 7,220 and the level-2 sample of 6,860. Thus, prior to merge with the level-1 data, weighted results from the level-2 sample represent the population of schools with greater than 75 students. Upon the school-level merge with the teacher-level data, the school-level drops an additional 4 percent (272 cases) to 6,580.

The cause of this change is teacher nonresponse. Again, the final teacher weight accounts for school- and teacher-level nonresponse. Thus, I did not perform missing data analysis.

Table 4

School-Level Missing Data Analysis

	Sample					
	Base (N=7,220)		Level 2 (N=6,860)		Principal (N=360)	
School characteristics	Mean	se	Mean	se	Mean	se
Enrollment (number of students)	567.90	7.08	564.02	7.13	647.97	31.27
Minority student population	39.48	0.70	38.71	0.69	55.26	3.41
Free- and reduced-price lunch population	45.96	0.64	45.55	0.63	54.52	2.83

Teacher-level changes occurred as a result of the school-level sample modifications. First, the sample dropped by 9.4 percent to 34,630 teachers as a result of the restriction to schools with at least 75 students. Second, I merged the teacher sample with the principal sample to identify how many teacher observations would be lost based on principal nonresponse. The result of the merge was a loss of 1 percent of teacher observations to 34,190 teachers. Although the loss in some cases was due to principal nonresponse, for which the final teacher weight does not account, I did not perform missing data analysis because of the relatively small change. Finally, due to teacher nonresponse the teacher sample lost an additional 3 percent of cases when merged with the school-level data (principal and school data included only schools and principals with corresponding data). Again, the final teacher weight accounts for teacher nonresponse so I did not perform any missing data analysis. The 2007–08 analytic samples contain 33,110 teachers and 6,580 schools.²² In the next few pages, I describe four types of

²² I do not round sample sizes in the analytic data set.

variables for this proposed study: outcome variable, independent variables, control variables, and weighting variables.

Outcome variable. The outcome, or dependent variable, is teacher-perceived autonomy. The empirical studies reviewed above often employed data reduction techniques to capture the construct of teacher-perceived autonomy. One purpose of data reduction techniques is to discover latent constructs in data that other statistical methods cannot measure directly. To that end, I created a factor and summative scale to represent teacher-perceived autonomy from responses to the six-part SASS question “How much control do you feel you have IN YOUR CLASSROOM over each of the following areas of your planning and teaching: selecting textbooks and other instructional materials; selecting content, topics, and skills to be taught; selecting teaching techniques; evaluating and grading students; disciplining students; determining the amount of homework to be assigned.”²³ Each of the four waves of SASS data included this question with no variation in the wording or order of the question. Respondents indicated their level of agreement with each statement using a Likert scale ranging from “No Control” to “Complete Control.” Table 5 lists the descriptive statistics for each of the six variables over time.

Based on changes in the SASS instruments from one administration to the next, the structure of the underlying measures required manipulation prior to incorporation into the outcome variable. Although the wording of the items did not change, the response options for the items did change. Specifically, the Likert scale used in years 1993–94 and 1999–2000 differed from one another and from those used in 2003–04 and 2007–08. In

²³ Emphasis (capital letters) included in original text.

1993–94 the scale was 0 to 5; in 1999–2000 it was 1 to 5; in both 2003–04 and 2007–08 it was 1 to 4.

Table 5

Variables Included in Teacher Autonomy Construct

Variable	1993 (N=44040)		1999 (N=38740)		2003 (N=40280)		2007 (N=33110)	
	Mea n	Std Dev	Mea n	Std Dev	Mea n	Std Dev	Mea n	Std Dev
select text	2.99	0.863	2.84	0.890	2.83	0.993	2.72	1.049
select content to be taught	3.10	0.793	2.89	0.866	2.95	0.938	2.70	1.043
teaching technique ^a	3.60	0.533	3.52	0.593	3.64	0.573	3.64	0.573
evaluate/grade students	3.61	0.522	3.57	0.548	3.69	0.530	3.55	0.624
decide discipline	3.29	0.697	3.22	0.725	3.51	0.530	3.47	0.664
assign homework	3.64	0.552	3.58	0.590	3.69	0.671	3.65	0.597

^a 2003 and 2007 unrounded mean values are 3.6402102 and 3.6444187, respectively; unrounded standard deviation values are 0.5728461 and 0.5729188, respectively.

I used a two-step process to rescale the responses from the first two SASS waves such that all four waves had a common Likert scale. First, I added a constant (1) to the scale used in 1993–94 to avoid complications that may arise from having a score of zero. Next, I converted the scales in 1993–94 and 1999–2000 to range from 1 to 4. I did not collapse the values but maintained the distinction of six and five data points, respectively, to represent the data within each year.

The cost of this transformation is that the scales' intervals differ, such that the difference between one response and the next (for example, a 1 compared to a 2) are not uniform across years. To illustrate, any teacher who responded 4 in 1993, with the intent of selecting one full unit lower than the “complete control” category, converted to a value of 3.4, which is only 0.6 units below “complete control.” Likewise, a teacher who

responded 4 in 1999–00, converted to a value of 3.2; whereas, a teacher who responded 4 in 2003–04 or 2007–08 continued with a value of 4. Thus, any direct comparisons of values from 1993–94, 1999–2000, and 2003–04 require cautious interpretation; whereas, direct comparisons are possible between 2003–04 and 2007–08. However, given the sample size available through SASS, such as the coefficient of variation, are robust to the metric of the six variables that form the outcome variable in this study, and are useful to identify patterns that exist across data points. Subsequent to the Likert-range transformation, for each year of data I used data reduction techniques to create standardized summative scales and factor scores, discussed in turn below.

Scale. A summative scale is just that – a nonweighted sum of the scores on the constituent measures that create a continuous variable. The basic assumption with this form of scale creation is that the aggregation of the elements included in the scale approximate a linear relationship with a latent construct. In this study, the scale produces a score for an individual teacher that is the sum of the teacher’s scores on each of the six items. For instance, the scale returns a score of 6 for a teacher that selected a 1 (or its equivalent, given rescale from 0 to 1 for 1993–94 data) on all six questions. Conversely, a teacher that selected a 4 or its equivalent, for every item would receive a score of 24.

Statisticians often evaluate composite scales using Cronbach’s alpha to indicate the reliability or internal consistency of the scale. Effectively, the alpha adds to the validity of an interpretation of the scale as a proxy for the construct in question. What constitutes a “good” Cronbach alpha value depends on the use of the scale. A generally accepted rule is that an alpha above 0.70 signifies a scale that is appropriate for analysis, particularly in a social science study rather than for assessment purposes. Table 6

presents the alphas for each of the autonomy scales. The alphas range from 0.79 to 0.75, which indicates an appropriate scale for the analyses in this dissertation.

Table 6

Cronbach alphas

	1993 (N=44040)	1999 (N=38740)	2003 (N=40280)	2007 (N=33110)
Autonomy Scale	0.7933	0.7854	0.7793	0.7452

Beyond the Cronbach alpha, other characteristics of the autonomy scale and the data that constitute the scale are important. For example, ideally the nonweighted scale score would be normally distributed, but as one would expect based on the mean values of the underlying variables presented in table 6 above, the scale skews toward the higher values in the range of 6 to 24. It is this skew that created the need to investigate the ability of the summative score to represent a continuous measure of a normally distributed distribution. The next subsection details how I investigated the appropriateness of the scale as outcome variable in two ways. First, I created descriptive statistics that confirmed that the summative scale was indeed similar to the construct of the outcome variable produced through factor analysis. Second, I produced coefficients from 2007–08 data with a multinomial multilevel model that used the categorical dependent variable as the outcome variable.

Factor. Factor analysis is a set of powerful analytical processes that can lead to construction of a single variable or measure built from several variables that expresses a construct that is not directly measured (i.e., a latent construct). Ideally, factors are normally distributed and have a mean of zero and a standard deviation of 1, which is important because this normalization permits use of the factor in any number of statistical procedures that assume a normal distribution, including ordinary least squares regression and hierarchical linear modeling. Typically, factors have advantages over summative

scales because during factor creation the process weights each constituent element by the shared variance it possesses with other variables that make up the factor. This weighting is desirable for use in any form of analysis because it suggests less measurement error. For example, I constructed a factor based on six categorical variables that in theory would have less reliability individually than they would as a composite.

The factor analysis that I conducted provides important information for this study regarding the selection of outcome variable. Based on factor loadings, both the values and the orders, presented in Table A.1, coupled with the Eigenvalues and percentage variance explained, both presented in Table A.2, I can deduce the following:

1. Combined, the six measures indicate that teachers believe that they have broad control over aspects of the classroom domain;
2. The six measures contribute in a statistically meaningful way to the construct; and
3. The factor loadings indicate that the construct is consistent over time because the loadings have a consistent pattern across all years, with one exception in year 2003–04, between text and discipline.

Simply put, I interpret the factor to be a reasonable proxy for teacher perception of autonomy in the classroom.

In Table A.3, I present coefficients for a linear multilevel model that has the standardized z-score as the outcome variable, and a multinomial, multilevel model that has the categorical variable as the outcome variable. Both models use the same variables and the same 2007–08 data. The multinomial models indicate the odds of teachers belonging to the high autonomy group vs. the low group, and to the moderate autonomy group vs. the low group. Through comparison of the direction of the coefficients and the

level of significance between the continuous multilevel model results and the multinomial, multilevel model results for the “high” category, I was able to confirm congruence between the model results. Thus, I argue the scale is robust to the underlying skew and is suitable as the outcome variable in this study.

Indicator and control variables. Indicator and control variables control for countervailing explanations of an outcome. The primary relationship of interest in the 2007–08 multilevel regression models in this dissertation is the relationship between teacher perception of autonomy in the classroom and school-level AYP status. The models are able to examine this relationship because in 2007–08, principals reported whether their school made AYP status at the end of the 2006–07 school year. Further, if the school did not make AYP, principals answered a follow-up question asking whether the school was in “in need of improvement” status based on at least two consecutive years of AYP failure in the same content area. Discussed previously, the distinction between years 1 and 2 of AYP failure is important because year 2 triggers NCLB-required sanctions.

To account for influences on teacher-perceived autonomy and isolate my hypothesized relationship between AYP status and teacher autonomy, I include indicator and control variables in the models based on the body of research reviewed in Chapter 2. At the teacher level, I include a variable to indicate that a teacher’s main assignment primarily addressed tested or not-tested content, as determined by NCLB testing requirements, and the grade-level of students. I have categorized this variable into four categories determined by elementary or secondary teaching and tested or not-tested main assignments according to NCLB requirements. I also include a factor score for teacher

perception of school leadership (see Table A.4 for details on factor score) and two measures of experience level. The first experience measure is an indicator of novice status; the second is a continuous measure of years as a teacher. As teacher-level control variables, I include demographic variables (gender and race). School-level indicator variables include AYP status for 2006–07, Needs Improvement status for 2007–08, and school-wide Title I status. School-level control variables include a three-category school-level indicator and continuous measures of total enrollment, percentage of minority students, and percentage of students approved for free- and reduced-price lunch.

I account for state-level sanction policies that are a response to federal accountability with two additional variables: sanction and non-sanction. I created these variables based on the state-level policy to either apply sanctions to any school that does not make AYP or to apply sanctions only to those schools that receive Title I funding. Rather than use these variables as indicators in the teacher- or school-level models, I use them to split the full sample into a sanction subsample and a non-sanction subsample. This approach maintains the integrity of the state-level variable and permits me to contrast the school and teacher characteristics that fall within the subsamples. Table A.5 identifies the states for each group.

Weighting variables. In addition to modeled variables, SASS data sets include replicate- and final-weight variables to adjust for the complex sample design, probabilities of sample selection, and subgroup nonresponse. The final teacher weight adjusts for the probability of selection at both the teacher and school levels. Specifically, the teacher-final weight is “the product of (Initial Basic Weight) and (School Sampling Adjustment Factor) and (Teacher List Nonresponse Adjustment Factor) and (Teacher-

Within-School Nonresponse Adjustment Factor) and (First-Stage Ratio-Adjustment Factor) and (Teacher Adjustment Factor)” (Tourkin et al., 2010, pp. 191-192). The HLM analyses use the final teacher weight variable²⁴ to permit generalizability of findings to the U.S. population of public school elementary and secondary teachers. To interpret coefficients as effect sizes and to model fixed and random effects appropriately in the multilevel models, all measures in the statistical analyses are standardized z-scores, grand-mean centered, or group-mean centered. Specifically, all continuous measures are standardized z-scores. Further, all categorical variables except novice status are grand-mean centered to reflect the average proportion of each characteristic across the sample (i.e., overall average). Novice and quality of leadership (z-scored continuous variable) are group-mean centered to reflect the average across the group (i.e., school-level average). Table 7 lists all variables and their respective descriptions.

²⁴ Tchwtg for 1993–94 and tfnlwgt and for other years.

Table 7

Description of Variables Used

Type	Name	Description
DEPENDENT VARIABLES:		
Scale	Teacher perception of autonomy in the classroom	Standardized scale calculated from six teacher-level measures (perception of control over: selecting class materials, selecting class content, selecting teaching technique, evaluating students, deciding discipline, assigning homework)
Tertile	Teacher perception of autonomy in the classroom	Categorical variable to indicate high/moderate/low levels of perceived autonomy created from the factor derived from the same six teacher-level measures listed above; used only in confirmatory analysis provided in Appendix.
INDICATOR VARIABLES:		
Teacher level	Perception of leadership quality [LEADER]	Continuous variable aggregated from factor of three Likert scaled variables (1 – strongly agree; 4 strongly disagree) teacher level variables: Administration is supportive, Principal enforces school rules and supports teachers enforcement, Principal communicates with staff,
	Years of experience [EXP]	Continuous variable measuring total years of teaching experience
	Novice status – less than or equal to 3 years experience [NOVICE]	Categorical variable where 0 = Novice; 1 = Non-novice
	Elementary – tested [GRADE_TEST]	Categorical variable where 0 = teaching level is elementary and main teaching assignment is general elementary, mathematics, natural sciences, English/language arts, or special education
	Elementary – not tested [GRADE_TEST]	Categorical variable where 1 = teaching level is elementary and main teaching assignment is other than those included in GRADE_TEST=0
	Secondary – tested [GRADE_TEST]	Categorical variable where 2 = teaching level is secondary and main teaching assignment is mathematics, natural sciences, English/language arts, or special education
	Secondary – not tested [GRADE_TEST]	Categorical variable where 3 = teaching level is secondary and main teaching assignment is other than included in GRADE_TEST=3

Table 7

Description of Variables Used (continue)

Type	Name	Description
DEPENDENT VARIABLES:		
School level	AYP status – Year 1 Failure [S_FAIL]	Dichotomous variable where 0=Yes and 1=No
	AYP status – Years 2 or beyond failure [NDIMP]	Dichotomous variable where 0=Yes and 1=No
	Title I status [TI_SW]	Dichotomous variable where 0 = Title I school; 1 = not Title I
	State applies sanctions to all schools that fail AYP [SANCTION]	Dichotomous variable where 1=Yes and 0=No (Source: NCES)
	State applies sanctions to Title I schools only [NO_SANCTION]	Dichotomous variable where 1=Yes and 0=No; (Source: NCES)
CONTROL VARIABLES:		
Teacher level	Sex [FEMALE]	Categorical variable where 0 = female; 1=male
	Race [RACE]	Categorical variable where 0 = non-white; 1=white
School level	School enrollment [ENROLL]	Continuous variable measuring total enrollment
	School % Minority Students [MINENR]	Continuous variable measuring the percentage of minority students
	School % Students FRPL [FARMS]	Continuous variable measuring the percentage of students approved for free- and reduced-price lunch
	School level[LEVEL]	Categorical variable where 0 = Elementary school (grades k-6); 1 = Combined school (grades k-12; 2 = Secondary school (grades 7-12)

Statistical Procedures

In this dissertation I utilize descriptive and multilevel analysis techniques. The analyses have two perspectives. The descriptive analysis of the autonomy scale takes a broad perspective of change from 1993 to 2007, whereas the HLM analyses yield a specific portrait of school and teacher characteristics that associate with teachers' perceptions of autonomy in the classroom for 2007–08.

Analyses of variation. To answer the over-time component in research question 1a, I offer three statistical measures of the autonomy scale for each year of data: means, coefficients of variation (CV), and intra-class correlation coefficients (ICC). The arithmetic mean is a measure of central tendency that is influenced by the metric of the underlying included variables. The inconsistent metrics of the constituent measures limit my ability to compare mean values over time. However, the mean is an important starting point for the discussion on teacher-perceived autonomy because it orients the findings to the average level of autonomy perceived by teachers in a given year. In contrast to mean values, the CV is a statistic that indicates variation around the mean. The CV is useful to compare values between data sets that have different units or metrics and is calculated as the standard deviation divided by the absolute value of the mean:

Formula 4.1
$$c_v = \frac{\sigma}{|\mu|}$$

Because the CV indicates total variation around the mean, it does not distinguish between variance attributable to teacher characteristics and variance attributable to school characteristics. Thus, I develop a fully unconditional model for each wave of data to identify the proportion of the total variation in my dependent measure that may be attributed to schools (Raudenbush & Bryk; (2002). These unconditional, multilevel

models do not have any indicator or control variables. They produce measures of variation in the teacher perception of autonomy (level-1) that is attributable to schools (level-2). The teacher-level model is

Formula 4.2
$$Y_{ij} = \beta_{0j} + r_{0j}$$

where Y_{ij} is teacher perception of autonomy for teacher i in school j , β_{0j} is the mean teacher perception of autonomy in school j , and r_{0j} is the random teacher effect or error term, which I assume to be normally distributed. Similarly, the unconditional model at the school level is

Formula 4.3
$$\beta_{0j} = \gamma_{00} + u_{0j}$$

where γ_{00} is the grand mean of teacher perception of autonomy for all schools, and u_{0j} is a random effect associated with unit j . I assume the school variation in teacher perception of autonomy has a normal distribution. The combination of the models for the two levels results in the fully unconditional model:

Formula 4.4
$$Y_{ij} = \gamma_{00} + u_{0j} + r_{0j}$$

The fully unconditional model partitions the total variance in the dependent variable into teacher- and school-level estimates. The equation for the ICC is

Formula 4.5
$$ICC = \tau_{00} / (\sigma^2 + \tau_{00})$$

where τ_{00} is the estimated between-school variance, and σ^2 is the estimated within-school variance for teacher perception of autonomy. This partition is integral to the over-time analyses because it indicates whether schools have become more integral over time to any variation in teacher-perceived autonomy. Simply put, the ICCs indicate how much

schools matter to teacher perceptions of autonomy. An examination of the ICCs over time can indicate whether the importance of schools to teacher-perceived autonomy increases or decreases over time. For, example, an increasing ICC from survey year to survey year would indicate that teachers in a particular school have become more similar to one another in their perceptions of autonomy. That is, their school grouping influences their perceptions when compared to teachers in another school. It may be that teachers in the same school react differently to the policy (e.g., teachers in NCLB-tested subjects vs. subjects not tested), or it may be that average differences between teachers in some schools are vastly more pronounced than those in another school (e.g., Title I schools compared to others); or it may be some combination of these and other school factors.

ICC comparisons. ICCs are dependent on the data sample from which they are drawn. Thus, to compare one ICC to another requires an estimate of the bias and standard error of the ICCs for each time period. In this dissertation, I used jackknifing to produce standard errors and confidence intervals for the ICC in each year. Jackknifing is a common procedure to measure accuracy of statistical estimates, particularly with non-normally distributed data (Barker, 2005). Essentially, jackknifing involves drawing a large number of random subsamples from the larger dataset to replicate the HLM analyses that yield the ICC. The more replications, the more accurate the reliability estimates produced from the jackknifed datasets. To compare estimates from two or more datasets, I completed the jackknife procedure for each year of data.

For this study, I created jackknife samples from the analytic dataset by randomly removing eight schools (approximately $1/1000^{\text{th}}$ of the base school-level sample) from the analytic sample without replacement. For each of these 1,000 jackknife samples in

each year, I ran fully unconditional models and calculated 1000 ICC estimates. Based on these 1,000 ICC estimates, I was able to calculate confidence intervals and standard error estimates for the ICC. Therefore, for each year of data in this study, I have ICC confidence intervals from 1,000 jackknife samples, which I can then compare from one dataset (or year) to another.

Hierarchical Linear Model (HLM) analysis. To answer the point-in-time components of my research questions and to investigate the effect of teacher and school-level measures, specifically AYP status, on teacher-perceived autonomy in 2007–08, I create a multilevel model based on Raudenbush and Bryk (2002) HLM approach. For 2007–08 data, in addition to the unconditional model, I develop a within-school model (level 1) and a between-school model (level 2), and use indicator and control variables to model the perception of autonomy for teacher i in school j . For the level-one or teacher level model, I examine perceived autonomy as a function of teacher assignment, experience level, gender, race, and perception of leadership quality. The level 1 model is

Formula 4.6
$$Y_{ij} = \beta_{0j} + \beta_{1j}(\text{GRADE_TESTED}) + \beta_{2j}(\text{NOVICE}) + \beta_{3j}(\text{EXP}) + \beta_{4j}(\text{FEMALE}) + \beta_{5j}(\text{RACE}) + \beta_{6j}(\text{LEADER}) + r_{0j}$$

where β_{0j} is the school mean of teacher perception of autonomy in school j ; β_{1j} is the mean difference between teachers based on assignment, categorized as those who teach elementary or secondary students in tested or non-tested subjects, respectively; β_{2j} is the mean difference between teachers who have less than or equal to three years of experience and those who have more than three years of experience; β_{3j} is the mean difference between non-novice teachers based on years of experience; β_{4j} is the mean difference between male and female teachers; β_{5j} is the mean difference between white

teachers and teachers of all other race/ethnicity categories; β_{6j} is the difference between teachers based on their perception of school leadership quality; and r_{0j} is the residual or error term for teacher i in school j .

I also model the associations between school characteristics and teacher-perceived autonomy at the school level. The level 2 or school-level model is

Formula 4.7
$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{AYP}) + \gamma_{02}(\text{NDIMP}) + \gamma_{03}(\text{T1}) + \gamma_{04}(\text{LEVEL}) + \gamma_{05}(\text{ENROLL}) + \gamma_{06}(\text{FARMS}\%) + \gamma_{07}(\text{FRPL}\%) + \gamma_{08}(\text{MINENR}\%) + \gamma_{09}(\text{AGGLDR}) + u_{0j}$$

where γ_{00} is the school average of teacher perception of autonomy; γ_{01} is the average difference between teacher perception of autonomy in schools that made AYP and in schools that did not make AYP for the first year; γ_{02} is the average difference between teacher perception of autonomy in schools that made AYP and in schools that did not make AYP for two consecutive years; γ_{03} is the average difference in teachers' perception of autonomy associated with schoolwide Title I status; γ_{04} is the average difference in teachers' perceptions of autonomy between elementary, combined, and secondary schools; γ_{05} is the average difference in teachers' perceptions of autonomy based on total enrollment ; γ_{06} is the average difference in the perception of autonomy for teachers associated with a percentage change in the average school-level population of students eligible for free- and reduced-price lunch; γ_{07} is the average difference in the teacher perception of autonomy associated with a percent difference in the average school-level minority student population; γ_{08} is the average difference in teacher perception of autonomy based on the average perception of school-level leadership quality; u_{0j} is the school-level residual or error term for school j .

My goal is to disentangle the teacher and school characteristics that comprise the variance in teacher-perceived autonomy. I have grand-mean centered all variables except novice status and teacher-level perception of school leadership quality. I model these variables with random effects and group-mean center both variables. I have also added an interaction term between novice status and school level to investigate whether the association between novice status and perceived autonomy varies by elementary, combined, and secondary school levels.

In addition to the models for the overall analytic sample, I divide the sample into two subsamples to examine teachers in states that apply sanctions to all failed AYP schools apart from teachers in states that follow NCLB and apply sanctions only to schools that receive Title I funding. This split sample permits me to compare teachers based on state policies on sanctions and to compare each of the split samples with the overall sample. Thus, I am able to capture school-level and teacher-level effects on teacher perceptions of degree and scope of autonomy within specific state contexts related to SBA reforms.

Limitations of the Study

This dissertation provides valuable insight into teacher-perceived autonomy in an SBA reform context. However, the study has six limitations of data and interpretation that merit the reader's attention. First, the analyses presented here do not include data collected through qualitative investigation techniques. The teacher "voices" that inform the study are confined to categorical responses that may or may not capture the full meaning of teachers' beliefs about autonomy in the classroom. To overcome this limitation, I have attempted to bolster the quantitative approach of this study with

findings that span time periods and education settings from scholars who produced mixed-methods studies and qualitative research. Second, although not unique to this study, a noteworthy limitation is that researchers cannot directly measure teacher autonomy (Allensworth et al., 2009; Ingersoll, 1996; Ingersoll & May, 2010, 2011). The construct of teacher autonomy is a proxy measure of a nuanced notion rather than a direct measure of an observable characteristic. Further, the scope of the construct is quite narrow as it specifies teacher-perceived autonomy in the classroom domain, rather than a more robust measure that might capture the scope of autonomy across curricula, budgets, and staffing; or that might capture additional domains of autonomy, such as districts, schools, and classrooms.

The third limitation is that the study contains limited AYP status data. Specifically, SASS data provides AYP status only in the 2007–08 data. Although 2003–04 data collection was post-NCLB implementation, NCES did not introduce AYP status into the surveys until 2007–08. Therefore, the 1993–94, 1999–2000, and 2003–04 data offer an assessment of teacher-perceived autonomy under SBA reform prior to NCLB, and the 2007–08 data serve an examination of teacher-perceived autonomy under the federal SBA reform, aka NCLB, roughly five years after the Act’s implementation. It may have served the study to have had two years of SBA reform prior to NCLB and two years post-NCLB for a stronger pre- and post-picture specific to NCLB. However, as it is, the study offers compelling findings regarding SBA reform. Fourth, although several scholars (notably Ingersoll) use SASS data for over-time comparisons, NCES does not specifically intend scholars to use the data for longitudinal comparisons. There is no panel weight available and questionnaires do not repeat all questions from the previous

data collection period. Further, as demonstrated by this study, the changes in Likert scales for those questions that do carry over from survey year to survey year complicate the ability to make direct comparisons over time. However, descriptive statistical analyses reveal consistency between the autonomy construct from year to year and reveal a pattern of change in teacher-perceived over time. Thus, this study does not attempt to use SASS beyond its capabilities.

The fifth limitation is that SASS is a self-administered, pencil and paper data collection process. Although school administrators and principals have detailed instructions to aid in survey completion, self-reported data are always subject to misinterpretation and misinformation (Boyd, et al., 2011; Ingersoll & May, 2011). The sixth and final limitation is that SASS data are not experimental data. Hence, the relationships discussed cannot be determined to be casual relationships. The statistical models do not capture all intervening effects on the outcome measure because quasi-experimental research does not offer the same level of control of treatment groups or random assignment as would an experimental design. Thus, readers should interpret the outcome measure with caution.

Chapter 4: Results

This chapter consists of three sections organized according to the research questions presented in Chapter 1. The first section offers descriptive results from the over-time analyses that address question 1a: From 1993 to 2007, how has degree of teacher-perceived autonomy in the classroom changed? The second section presents descriptive statistics and HLM findings that address question 1b: For 2007–08, what school and teacher characteristics are associated with differences in teacher-perceived autonomy?; question 2a: How do teachers who work in schools that made AYP perceive their autonomy compared to those who work in schools that did not make AYP?; and question 2b: How do teachers who work in schools in the first year of failed AYP status differ from those who work schools in the second year or later of failed AYP status. Section three concludes this chapter by extending analyses that address research question 1b to consider state-level policy regarding AYP sanctions and Title I status. In total, these three sections provide insight into answering the motivating question of this study: How do teachers perceive their own autonomy in the classroom within the policy context of SBA reforms, specifically NCLB? The chapter closes with a summary of the results.

Over-Time Analyses: 1993–94, 1999–2000, 2003–04 and 2007–08

To examine changes in teacher-perceived autonomy over the last two decades, I used the four data points of SASS data between 1993–94 and 2007–08. As described in chapter 3, I transformed the six Likert-scaled variables contained in the final autonomy measure such that each of the six variables ranged from a low score of 1 to a high score of 4. However, I kept the number of data points of the Likert scale for each year (e.g., responses from 1993–94 still contain 6 measurement points with the lowest point of 1 and

the highest of 4). Specifically, 1993–94 has 6 points, 1999 has 5 points, and each of the two later years have 4 points.²⁵ The advantage of this approach is that each year has the same range of values that correspond at the low and high values. The disadvantage is that the underlying scales continue to present a challenge for direct comparison of means of the autonomy measure from year to year.

To illustrate, Figure 4 presents the weighted mean values for the autonomy scale across all teachers in the sample for the specified year.

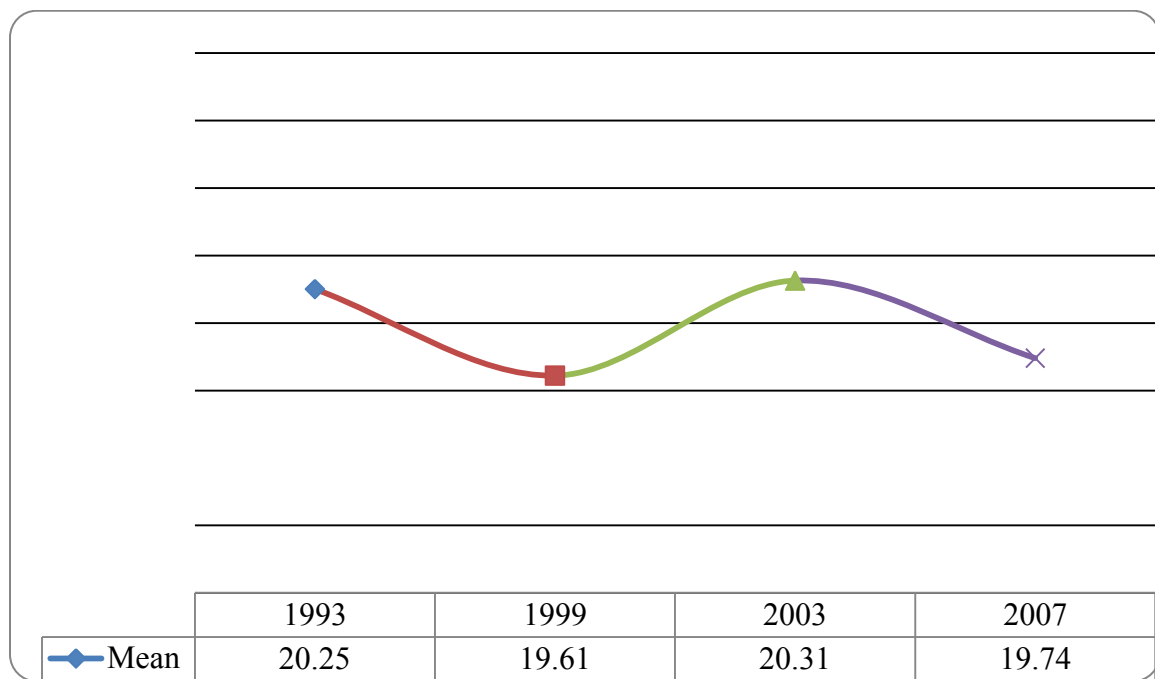


Figure 4. Autonomy scale mean values across teachers: 1993–94, 1999–00, 2003–04, 2007–08

The weighted mean values for the autonomy scale ranged from 19.6 to 20.3. Due to the properties of the scale, which differ between 1993–94, 1999–2000, and 2003–04, it is difficult to make conclusive comparisons between the mean values of autonomy across

²⁵ 1993–04 scale is 1.0, 1.6, 2.2, 2.8, 3.4, 4.0; 1999–2000 scale is 1.0, 1.75, 2.5, 3.25; 2003–04 and 2007–08 are 1.0, 2.0, 3.0, 4.0.

these three years. Indeed, the change from the first 2 data points and 2003–04 contradicts a pattern suggested by the change between 1993–94 and 1999–2000. While inconclusive, it is reasonable to postulate that the change that NCES made to the underlying Likert scale in 2003–04 that decreased the response values from 5 to 4 may partially explain the 2003–04 change. The change in 2003–04 may also relate to local conditions or an influx of resources; however, during the review of literature for this study, no evidence emerged as a systematic, national influence that would boost teacher responses to the SASS items for the 2003–04 academic year as opposed to any other year.

Given that the maximum value on the scale for each year was 24, the fact that the average in each year hovers near 20 suggests that teachers perceived a high average level of autonomy at each of the time points. In contrast to prior years, the shared underlying metric of the autonomy scale for 2003–04 and 2007–08 permits direct comparisons across these two waves. An independent *t* test reveals that the weighted average scores for each year (20.3 for 2003–04 and 19.7 for 2007–08) are statistically different from one another. Although still relatively high, the average level of autonomy in 2007–08 is lower than the average level in 2003–04, by about one-tenth of a standard deviation, which suggests a slight decrease in the teacher-perceived autonomy.

While differences underlying the metric of the autonomy scale across years create complications for comparison of absolute statistics like the mean or standard deviations, an alternative statistic, such as the coefficient of variation (CV), is more robust to these shortcomings in the data. To that end, Figure 5 depicts the amount of variation in the average autonomy score from one wave of data to the next.²⁶ The CV analysis indicates an upward trend in variation across the 4 years (from 1.06 in 1993 to 1.53 in 2007).

²⁶ Please see Table A.7: CV and ICC calculations, which lists all values used to calculate the CVs.

Taken in combination with the relatively high mean values on the autonomy scale, the larger CVs suggest that although teachers in general continue to perceive high levels of autonomy in the classroom, there is growing variability around that perception.

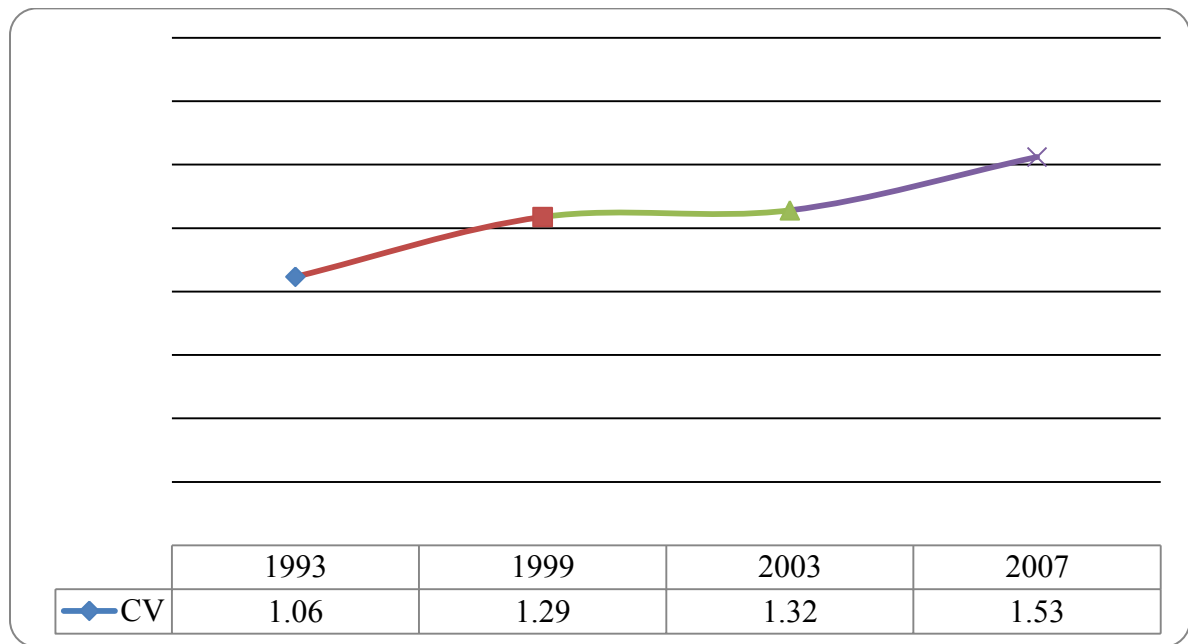


Figure 5. Coefficients of variation: 1993–94, 1999–2000, 2003–04, 2007–08.

Figure 6 shows the ICC values for each year.²⁷ Like the CV, the ICC is a measure of variability but it differs from the CV because it is a multilevel statistic. The ICC partitions the variance in a measure within and between groups. An ICC value of 1 indicates that all of the variability in a measure is between groups. An ICC value of zero indicates that none of the variability is between groups. The ICC analysis is important in the current study because it dissects the source of change in teacher-perceived autonomy between teachers and schools. In other words, the larger an ICC, the more teachers' perceptions of autonomy differ between schools, which hints at a single or multiple

²⁷ Table A.7: CV and ICC calculations lists all values used to calculate the overall ICCs.

school-level characteristic that influences teacher perceptions The ICCs reveal that school context has an increasing and measureable role in teacher-perceived autonomy. To compare ICCs statistically across years, I calculated confidence intervals for each ICC, and compared the values from each wave of data to determine if the confidence intervals overlapped.²⁸ Based on the confidence intervals, I was able to determine that the ICCs from each wave of data were statistically different from one another. ICCs used in this dissertation and their corresponding confidence intervals are located in Table A.6.

²⁸Please see the ICC Comparisons section in Chapter 3 for details on the confidence interval calculations.

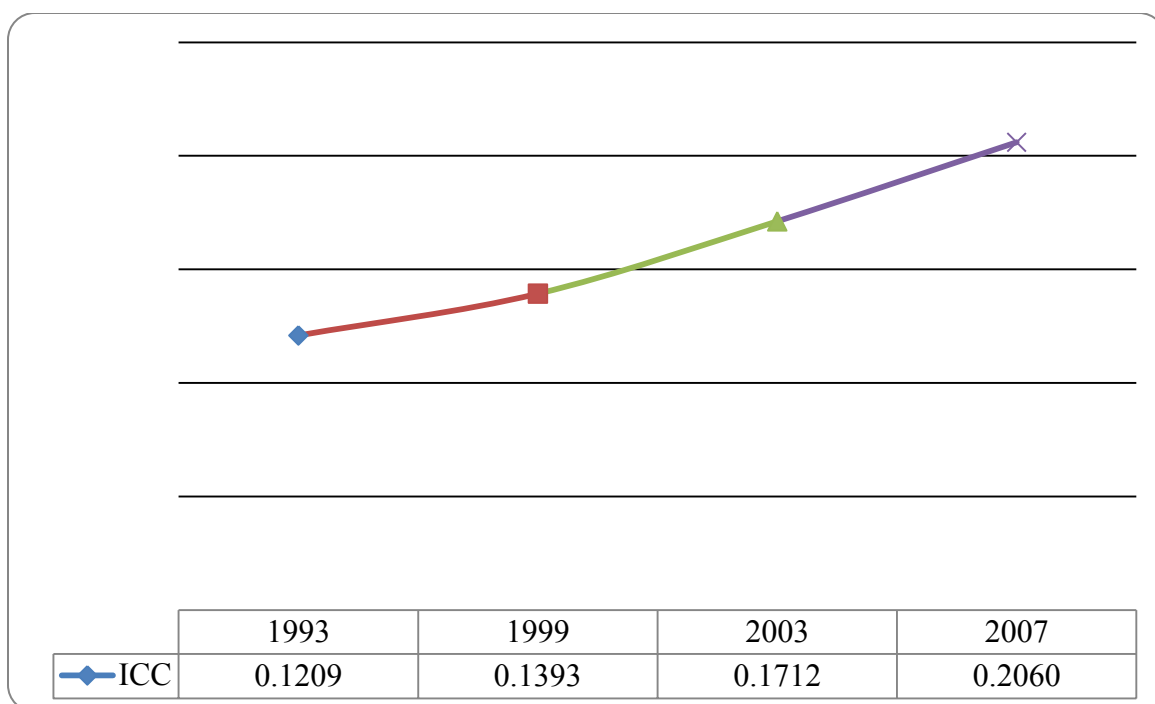


Figure 6. ICC over time: 1999–93, 1999–00, 2003–04, 2007–08.

Results from this analysis indicate that from 1993–94 to 2007–08, the import of school context on teacher-perceived autonomy increased substantially. Specifically, in 1993–94, about 12 percent of the variation in teacher-perceived autonomy was attributable to the school level. Over time, the school level portion of overall variation increased steadily such that in 2007–08, it reached nearly 21 percent. Effectively, the role of school context on teacher-perceived autonomy increased substantially over the time period by about 72 percent.

Given the ICC findings across the entire sample, I investigated the change in teacher-perceived autonomy for two subsamples of interest: teachers who work in Title I schools and teachers who do not work in Title I schools. As part of the investigation into the hypotheses regarding different teachers perceiving changes in autonomy differently, I examined whether there was more between-school variation in Title I schools compared

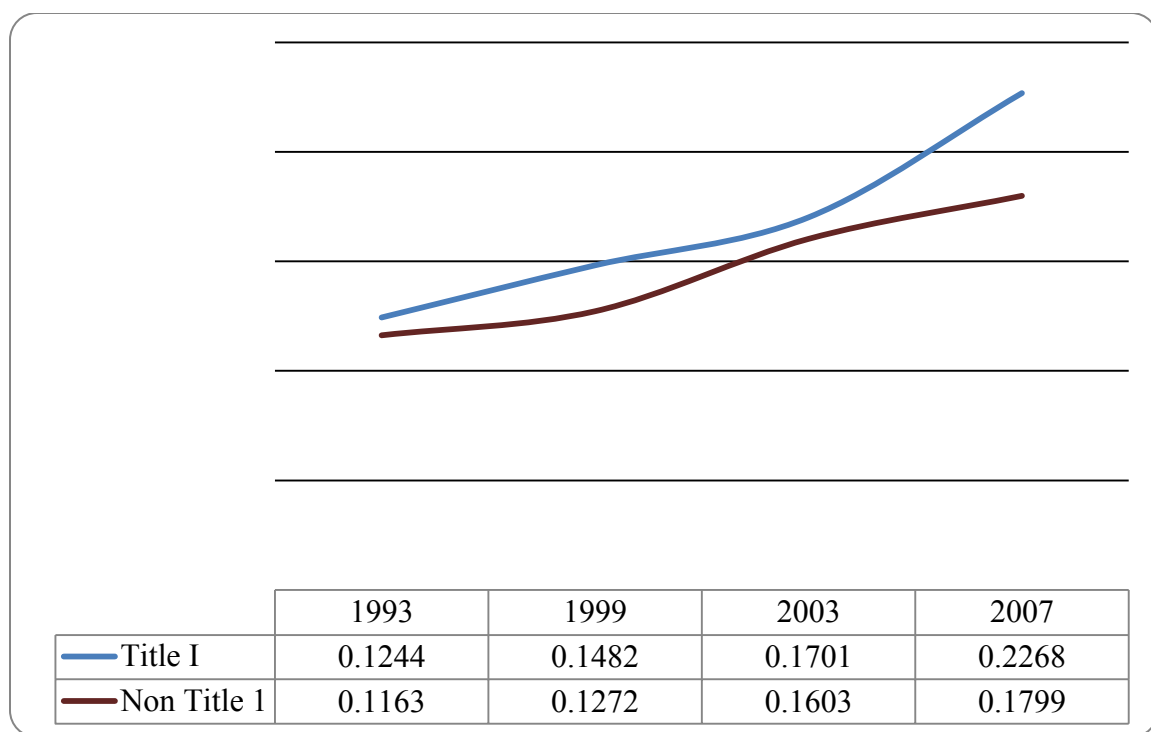


Figure 7. ICC over time by Title I status: 1993–94, 1999–00, 2003–04, 2007–08.

to non-Title I schools. Figure 7 illustrates the school-level variance of teacher-perceived autonomy by Title I status for each of the four waves of data. As with the overall ICCs, I generated confidence intervals for the Title I and non-Title I ICCs to compare the values over time and between groupings.²⁹ These comparisons afford three important observations that lead to the next set of analyses in this study. First, for each year the ICC for teachers at Title I schools is statistically different from the ICC for teachers at non-Title I schools. For example, in 1999 a greater portion of the total variation in teacher-perceived autonomy among teachers who worked in Title I schools was due to their school context (0.1482) compared to teachers in non-Title schools (0.1272). In other

²⁹ Please see Table A.8: CV and ICC calculations by Title I status, 1993-2007 for calculations.

words, compared to teachers in non-Title I schools, there was greater variability in teacher perceptions of autonomy between Title I schools.

Second, for both Title I and non-Title I schools for each succeeding wave of data, the ICC is statistically higher than the one before. Most notably, the 2007 ICC is statistically higher than the 1993 ICC. Thus, by 2007 the role of school context in teacher-perceived autonomy had increased for teachers in Title I schools, as it had for teachers overall. Specifically, from 1999 to 2007, for teachers in Title I schools the proportion of total variation in teacher-perceived autonomy attributable to their school context nearly doubled (82 percent increase). For teachers in non-Title I schools, in 1993 school-level characteristics explained 11.6 percent of the variation in average teacher-perceived autonomy; whereas in 2007, school-level characteristics explained 18.0 percent of the variation. This is an increase of 54 percent for teachers in non-Title I schools. In sum, the school context has come to matter more in teacher perceptions of autonomy overall; however, the effect is more pronounced for teachers in Title I schools

The third observation is that the spread in ICC values between teachers in non-Title I schools (0.1799) and those in Title I schools (0.2268) is largest in 2007–08. Thus, the school context for teachers in Title I schools appears to have gained importance over time in shaping teacher-perceived autonomy and is strongest in 2007–08, the year with the largest percentage of schools failing to make AYP (Usher, 2011). Taken together, the findings support a fuller investigation into the possible effects of SBA reforms, and more specifically NCLB, on teacher-perceived autonomy. The change in variability between Title I schools in average teacher-perceived autonomy, particularly between 2003-04 and 2007-08, suggest that changes in the policy context may be an important factor in

understanding teacher-perceived autonomy. The full sample analyses section below addresses whether teachers in Title I schools perceive lower or higher average levels of autonomy compared to their colleagues in non-Title schools.

As a whole, the descriptive analyses in this section indicate recent changes in the average levels of teacher-perceived autonomy, as well as increases in the variation around the average. Specifically, although teachers reported generally high levels in 2007–08, the average level of autonomy in 2007–08 was lower than that in 2003–04. Moreover, the amount of variation about the means has increased over time. Further, school-level factors or context appear to explain more of the total variation in autonomy over time, which indicates that school-level characteristics, practices, policies, or combination thereof may have more of a role influencing teacher-perceived autonomy in 2007–08 than in the 15 years prior.

The remaining sections present findings for the analyses of the 2007–08 data to offer evidence of change in teacher-perceived autonomy based on specific teacher and school characteristics. There are six subsections in the succeeding section. The first four subsections present results for the full analytic sample, and the subsequent two subsections present results for two subsamples of the 2007-08 data. I split the data into “sanction” states and “non-sanction” states based on state-level treatment of the federal requirement to apply sanctions to Title I schools that fail to make AYP.

Full Sample Analyses: 2007–08

Results from the descriptive analyses that examined teacher-perceived autonomy over the last two decades are important because they establish that teacher-perceived autonomy has not been static. Specifically, the descriptive results indicate that total

variability of teacher-perceived autonomy has increased over time and that the variation is increasingly between schools. However, these findings provide only a baseline description of changes and do not permit a more in-depth assessment of specific teacher-level or school-level characteristics that may account for differences in perceived autonomy. Such an analysis requires a more precise statistical method, like multilevel modeling. Although several characteristics are of interest in this study, of particular interest are the two indicators of AYP status: first-year failure and second-year or beyond failure. These indicators were not available prior to the 2007–08 SASS data file. Hence, the multilevel analyses are limited to the 2007–08 data.

Descriptive statistics. To ground the HLM findings, Table 8 presents the mean, standard deviation, and minimum and maximum values for the variables in the multilevel models. The weighted data represent a nationwide population of over 3 million teachers and nearly 91,000 schools. Just over 70 percent of the teachers in the weighted analytic sample teach “tested subjects” (e.g., English language arts, mathematics, and science), 44 percent of whom teach elementary students and 26 percent of whom teach secondary students. The average years of teaching experience in the sample is 14 years, with 17 percent of teachers who are novice teachers with only one, two, or three years of experience. Nearly 84 percent of teachers are white and 76 percent are female.

At the school level, 24 percent of schools failed to make AYP for at least one year and 12 percent failed to make AYP for a second consecutive year or more. Thirty-three percent of schools have school-wide Title I status; 72 percent are elementary schools, 6 percent are combined schools.

Table 8

Descriptive Statistics for the Analytic Sample: School Year 2007–08

(Weighted statistics; unweighted sample sizes: Teacher N = 33,110; School N = 6,580)

Teacher characteristics	Mean/ Frequency¹	SD	Min.	Max.
Assignment				
Elementary – tested	44.25	-	0	1
Elementary – not tested	7.55	-	0	1
Secondary - tested	26.16	-	0	1
Secondary - not tested	22.04	-	0	1
Experience				
Novice (less than or equal to 3 years)	16.74	-	0	1
More than 3 years	83.26	-	0	1
Sex				
Male	23.79	-	0	1
Female	76.21	-	0	1
Race				
All other races	16.12	-	0	1
White, non-Hispanic	83.88	-	0	1
Perception of leadership quality (factor score)	0.05	0.91	-3.17	0.93
Years of experience	13.66	10.46	1	52
School characteristics				
AYP Status – Year 1 failure	24.12	-	0	1
Needs Improvement Status – Year 2 failure	11.87	-	0	1
Title I status	33.37	-	0	1
School level				
Elementary	71.59	-	0	1
Combined	5.73	-	0	1
Secondary	22.68	-	0	1
Enrollment (number of students)	564.02	543.61	75	5000
Minority student population	38.71	32.16	0	100
Free- and reduced-price lunch population	45.55	26.30	0	100

¹Percentage unless otherwise stated.

schools, and 23 percent are secondary schools. The average total school enrollment size for the sample is 564 students. The average proportion of minority and free- and reduced-priced lunch eligible students in a school are 39 percent and 46 percent, respectively. In addition to descriptive statistics for each variable, given the prominence of Title I and AYP status in the current study, I determined the extent to which Title I schools and AYP

failure overlap in the data. Table 9 reports the intersection of these two statuses. The results indicate that 30 percent of Title I schools failed to make AYP as did 21 percent of non-Title I schools.

Table 9

Title I Status by AYP Status: School Year 2007–08

AYP Status	Title I status	
	Yes (%)	No (%)
Fail	29.51	21.23
Pass	70.49	78.77

The descriptive statistics are important because they situate the analytic sample. For example, the overwhelming majorities of teachers are white, female, teach tested subjects and work in elementary schools. To move beyond descriptive statistics and toward statistical relationships one must use regression. As discussed in chapter 3, given the nested nature of teachers and schools and the desire to tease out school effects from teacher effects, the appropriate type of regression is multilevel regression. The paragraphs that follow discuss results from one teacher-level model and four multilevel models, including two models that I developed from split samples based on state sanction policies for AYP status.

Multilevel model results. Table 10 presents three of five HLM models built with 2007–08 data that answer the remaining research questions: For 2007–08, what school and teacher characteristics are associated with differences in teacher-perceived autonomy?; How do teachers who work in schools that made AYP perceive their autonomy compared to those who work in schools that did not make AYP?; and How do teachers who work in schools in the first year of failed AYP status differ from those who

work in a school in the second year or later of failed AYP status? As a brief review, the conceptual underpinnings of the study include Dahl's (1984) constructs of degree, scope, and domain. *Degree* reflects the extent to which one party may influence another in the power relationship; *scope* defines the issue or content over which an individual, group, or entity may have power; and *domain* indicates over whom or over what systems actors may exercise power.

Table 10

Multilevel Models for Scope and Degree of Teacher-Perceived Autonomy in the Classroom Domain^a

	Model 1	Model 2	Model 3
Intercept	0.28 ***	0.40 ***	0.40 ***
Teacher characteristics			
Elementary_Test subjects	-0.69 ***	-0.56 ***	-0.56 ***
Elementary_Non-tested subjects	0.05	0.18 ***	0.18 ***
Secondary_Test	-0.39 ***	-0.36 ***	-0.36 ***
Female	0.00	0.01	0.01
Non-white	-0.10 ***	0.00	0.00
Perception of leadership ^b	0.26 ***	0.26 ***	0.26 ***
Novice (le 3 years)	-0.04 *	-0.07 **	-0.14 ***
Experience (years)	0.06 ***	0.06 ***	0.06 ***
School characteristics			
Fail AYP		-0.07 *	-0.07 *
Needs Improvement		-0.01	-0.01
Schl level - Elementary		-0.31 ***	-0.31 ***
Schl level - Combined		0.12 **	0.12 **
Schoolwide Title 1		-0.06 *	-0.07 *
Enrollment		-0.08 ***	-0.08 ***
% FRPL		0.00	0.00
% Minority		-0.13 ***	-0.13 ***
Ave schl perception of leader ^b		0.19 ***	0.19 ***
Interaction Terms			
Novice * level (Elem)			0.14 **
Novice * level (Combined)			-0.04
Random Effects			
Intercept	0.20 ***	0.18 ***	0.18 ***
Perception of leadership		0.18 ***	0.18 ***
Novice		0.32 ***	0.32 ***

* $p \leq .01$; ** $p \leq .001$; *** $p \leq .0001$

^a Coefficients reported in effect sizes. All continuous measures were standardized across the entire sample ($M=0$; $SD=1$).

^b The perception of leadership measure is a factor score based on four categorical measures. The average school perception of leadership is the leadership factor score aggregated to the school level.

I discuss the findings for each of the three models in turn below. To parallel the structure of the research questions, I discuss the findings related to AYP separately from the findings of the other teacher and school characteristics.

Teacher and school characteristics. Model 1 is the teacher-level model. This model indicates that without consideration of school characteristics, several teacher-level characteristics associate with teacher-perceived autonomy.³⁰

Model 1. On average, teachers that teach tested subjects perceive lower autonomy than those who teach non-tested subjects. Specifically, compared to teachers who teach secondary students in non-tested subjects (the reference group in the model), teachers who teach elementary students in tested subjects have a perception of autonomy that is lower by 69 percent of a standard deviation, and teachers who teach secondary students in tested subjects have a perception of autonomy that is lower by 39 percent of a standard deviation. In contrast, there is no significant difference in the perception of autonomy between teachers who teach elementary or secondary students in non-tested subjects. Simply put, testing matters to teacher-perceived autonomy, and it matters more for elementary teachers than it does for secondary teachers.

Also significant but smaller in magnitude are differences by race, perception of leadership, and experience levels. Compared to White teachers, non-White teachers perceive lower levels of autonomy in the classroom (by 10 percent of a standard deviation). As teacher perceptions of quality in their school leaders increase, so does their perception of autonomy (26 percent of a standard deviation). The more individual teachers perceive their principals as supportive, the more autonomy they perceive in the

³⁰ Beginning with model 1, all coefficients are standard deviations (or effect sizes). A single standard deviation change in a select characteristic equates to a percentage of a standard deviation change in teacher-perceived autonomy based on the coefficient of the select characteristic.

classroom. Meanwhile, teachers who have a standard deviation above the average years of experience (13.66 years) perceive higher levels of autonomy (by 6 percent of a standard deviation). In general, the more experience teachers have, the higher their perceptions of autonomy. However, there is a unique effect of being a novice teacher (a teacher with 3 years or fewer of experience). Novice teachers perceive lower autonomy than their non-novice peers by 4 percent of a standard deviation. This novice effect is over and above the experience effect. In other words, teachers with less experience have a lower perception of autonomy, but if the teacher is also a novice teacher, the difference is even greater.

As a whole, the teacher-level model indicates that testing and school leaders have substantive effects on teacher perceptions of autonomy in the classroom, and that experience and race play smaller but still statistically significant roles. I discuss novice findings in more detail in models 2 and 3 below, but the important takeaway from model 1 is that the effect on autonomy for novice teachers includes two components: an experience differential between non-novice and novice status; as well as an effect from novice status itself.

Model 2. The second model incorporates school-level variables into the analysis. Several school-level characteristics relate to teacher-perceived autonomy.³¹ On average and compared to secondary schools, teachers in elementary schools perceive 31 percent of a standard deviation less autonomy, and teachers in combined schools experience 12 percent of a standard deviation more autonomy. Teachers who work in Title I schools perceive 6 percent of a standard deviation lower autonomy compared to teachers who

³¹ This section includes all pertinent school-level characteristics apart from AYP status and Needs Improvement status. I present the findings for these two measures in the next section as question 2 discussions.

work in lower-poverty schools. Teachers in schools with above-average total enrollment and above-average proportions of minority students perceive lower autonomy compared to those who teach in schools with average total and minority enrollment (by 8 percent and 13 percent of a standard deviation, respectively). Meanwhile, positive effects associated with school-level variables exist for teaching in combined schools (compared to teaching in secondary schools) and in schools with an above-average schoolwide perception of leadership. In fact, even after controlling for the teacher-level perception of school leadership, teachers who work in a school with a collective perception of high-quality leadership experience a substantive and statistically significant increase in their perceived autonomy (19 percent of a standard deviation).

Compared to model 1, the school characteristics in model 2 alter several significance levels and the magnitudes for all but two³² teacher-level coefficients. The race category loses significance; the category of teachers who teach elementary students in non-tested subjects becomes significant; and the novice category changes significance to the 0.001 alpha level. Regarding coefficients, although tested subjects at both grade levels continue to have large negative coefficients, the magnitudes for elementary has fallen by 0.13SD, while that for elementary non-tested has increased by 0.13SD. Essentially, by controlling for the elementary school level, the negative 0.56SD coefficient is the effect what remains for teachers of elementary students in tested subjects and the positive 0.18SD is what remains for teachers of elementary students in non-tested subjects. Thus, of teachers who teach elementary students, those who teach in tested subjects perceive lower autonomy on average compared to teachers who teach any

³² Experience and perception of leadership

grade level in non-tested subjects by nearly 75 percent of a standard deviation.³³

Similarly, by 36 percent of a standard deviation, teachers of secondary students in tested subjects perceive lower autonomy compared to those who teach secondary students in non-tested subjects.

On its own the coefficient for novice teachers in model 2 equates to lower autonomy by 7 percent of a standard deviation compared to non-novice teachers, and experience remains at 6 percent of a standard deviation. The experience measure represents average years expressed in standard deviations of experience across the sample. This standardized measure for experience complicates interpretation of the novice effect because any standardized measure has a mean of zero. For the experience measure, the zero value corresponds to teachers who have 13.66 years of experience (the weighted average across the sample). To be clear, one can interpret the estimated regression coefficient for experience as 6 percent of a standard deviation increase in perceived autonomy for each standard deviation increase in teaching experience. However, this finding applies to teachers of average experience. By definition, novice teachers (teachers with only one, two, or three years of experience) are well below the average, which means they are below the zero mean of the standardized experience variable and have less than the 0.06SD effect size for experience. To get to the standard deviation value of experience for novice teachers requires identification of the mean values associated with each year of novice status.

Table 11 details how I arrived at the net effect of novice status for model 2 given the presence of a standardized experience measure in the model. A novice teacher with one year of experience would have a negative 1.23 standardized value for experience.

³³ $(-0.56SD + 0.18SD) = \text{absolute difference of } 0.74SD.$

The negative 1.23SD value converts the experience effect for first-year novice teachers from the positive 6 percent of a standard deviation to a *negative* 8 percent of a standard deviation.³⁴

Table 11

Experience Effect on Novice Status for Model 2

Experience Years	Experience SDs	Experience coefficient	Experience effect	+	Novice coefficient	=	Net effect
1	-1.23	0.06	-0.08		-0.07		-0.15
2	-1.14	0.06	-0.07		-0.07		-0.14
3	-1.04	0.06	-0.06		-0.07		-0.13
4	-0.95	0.06	-0.06		0.00		-0.06
14	0.00	0.06	0.00		0.00		0.00
25	1.06	0.06	0.07		0.00		0.07

To arrive at a net effect of experience and novice status of -0.15SD, the first-year teacher carries the -0.07SD effect from the novice coefficient in addition to the -0.08SD for experience. In other words compared to a non-novice teacher with average experience, a first-year teacher perceives lower autonomy by 15 percent of a standard deviation. Alternatively, a teacher who is one year past the novice stage (with four years of experience) still has a negative experience effect similar to that of novice teachers, but does not have an additional effect associated with novice status; however, a teacher who has 25 year experience is well above the average experience effect and has a positive experience effect of 7 percent of a standard deviation.

In addition to the fixed effects discussed above, model 2 includes random effects for novice status and perception of leadership. Both measures significantly vary across school settings. The random variation signifies that novice teachers' experience level

³⁴ $(-1.23\text{SD for novice experience} * 0.06\text{SD for average experience}) = -0.08\text{SD}$

affects teacher-perceived autonomy differently in different schools, as do teachers' perceptions of their school leaders. In sum, model 2 echoes the model 1 finding that teachers in tested subjects perceive lower autonomy compared to those who teach non-tested subjects and teachers with less experience, particularly novice teachers, perceive lower levels of autonomy compared to those with more experience. The random effects highlight a series of possible school-level effects (which I discuss later), including grade structure, school size, student enrollment characteristics, Title I and AYP statuses, and extend the positive teacher-level leadership effect to include an additional schoolwide leadership effect.

Model 3. The third model includes all measures from model 2 as well as an interaction term for novice by school level (novice by elementary and novice by combined) that attempts to account for some of the variability detected in the random effect for novice teachers.³⁵ The interaction term enables the model to separate the novice effect for teachers in elementary and combined schools from those in secondary schools. Coefficients in model 3 do not change from those in model 2 with the exception of the novice main effect (-0.07 SD to -0.14 SD) and the significant coefficient for the interaction of novice status and elementary level. Just as the experience measure in model 2 complicates the interpretation of the novice effect, the interaction in model 3 requires careful dissection to understand the overall effect for novice teachers by school level.

Without an interaction term, the novice main effect from model 3 would indicate that compared to teachers with average experience, novice teachers perceive lower

³⁵ The perception of leadership remains a random effect even though I do not model any interaction. Unfortunately, the models in this study do not capture the characteristic(s) that are the source of school-level variation. Nonetheless, the presence of the leadership measure in the random effects statement makes for a better fitting model.

autonomy by 14 percent of a standard deviation. Likewise, the school-level main effects would indicate that, regardless of experience, teachers who work in elementary schools perceive lower autonomy by 31 percent of a standard deviation compared to teachers who work in secondary schools, whereas those who work in combined schools perceive higher autonomy by 12 percent of a standard deviation. However, the addition of the interaction terms changes the interpretation of these main effects for novice and for school level. Effectively, the interaction term indicates that novice teachers who teach in elementary schools perceive autonomy no differently than do their more experienced peers. Thus, the novice effect occurs only for teachers who teach at the secondary level.

Table 12 details the novice effect for secondary teachers. To calculate the effect of novice teachers in secondary schools one must include the experience effect, the novice main effect, and the novice interaction effect. The first three columns of the table correspond to the experience effect for novice teachers and mimic Table 11 values.

Table 12

Experience and Interaction Effects on Novice Teachers in Secondary Schools for Model 3

Year	Average			Novice status	Novice coefficient	Net effect
	Experience (SD)	Experience coefficient	experience main effect			
1	-1.24	0.06	-0.08	1	-0.14	-0.21
2	-1.14	0.06	-0.07	1	-0.14	-0.21
3	-1.04	0.06	-0.06	1	-0.14	-0.20
4	-0.95	0.06	-0.06	0	-0.14	-0.06
14	0.00	0.06	0.00	0	-0.14	0.00
25	1.06	0.06	0.07	0	-0.14	0.07

The addition of the interaction term in model 3 increases the novice effect from negative 0.07SD to negative 0.14SD. However as the interaction term shows, this is neither a uniform effect for novice teachers, nor is it the final effect given the experience measure in the model. The experience effect for a first-year secondary teacher is -0.08SD and the main effect for novice status is -0.14SD. Combined, these effects indicate that, on average, a first-year novice teacher who works in a secondary school perceives lower autonomy by 21 percent of a standard deviation compared to secondary school teacher with average experience.

AYP status. Research question 2 asks how SBA reform, represented by a schools NCLB AYP status in 2007-08, associates with teacher-perceived autonomy in the classroom domain, and its sub-questions ask about the relationship specific to teacher-perceived autonomy and AYP status. In addition to the teacher and school characteristics discussed in the previous section, models 2 and 3 also include school-level variables for AYP failure in year one (Fail) and AYP failure for two consecutive years or more (Needs Improvement). As discussed earlier, two consecutive years of AYP failure trigger sanctions, which include the label of Needs Improvement for the school after the second year of failed AYP status. For both multilevel models, AYP status negatively relates to teacher-perceived autonomy. All else being equal, on average teachers who work in schools during the first year of AYP failure perceive lower autonomy by 7 percent of a standard deviation compared to those in schools that made AYP. In contrast, the perceived autonomy of teachers working in Needs Improvement schools was not statistically significantly different from the perceived autonomy of teachers in schools that met AYP. Perhaps surprisingly, AYP status matters in the first year of failure, but

teachers in schools that face sanctions due to at least two consecutive years of failure do not have different perceptions of autonomy compared to teachers in schools that made AYP.

Subsample analyses: 2007–08

To capture the AYP status and sanction relationship outlined in the empirical framework for the 2007-08 data and to incorporate state-level treatment of AYP sanctions, I split the sample according to state-level treatment of NCLB-required sanctions for schools that fail to make AYP (“sanction sample” and “non-sanction sample”). Thirty-two states³⁶ apply sanctions to all schools that fail to make AYP regardless of Title I status, which equates to a sanction sample of 78 percent of teachers nationwide; but only 19 states (22 percent of teachers nationwide) apply sanctions just to Title I schools that fail to make AYP. In the subsequent two sections I present descriptive statistics of the two subsamples, results from the multilevel regression model for each sample, and a discussion of how the regression results by state sanction policy differ from the full model and from one another.

Descriptive results. Table 13 lists the descriptive statistics for each subsample. Considering that the subsample for teachers in the sanction sample is more than three-quarters of the full sample, it is not surprising that few statistically significant differences³⁷ exist between the composition of the full analytic data and the sanction subsample. At the teacher level, the sanction subsample has slightly fewer White teachers and has teachers with slightly lower average experience compared to the analytic sample.

³⁶ Please see Table A.5 for details.

³⁷ I made comparisons using dependent t-tests for statistical significance at the $p < .05$.

At the school level the sanction subsample has nearly 11 percent higher average minority enrollment.

Table 13

*Descriptive Statistics for Subsamples by State Sanction Policies:
School Year 2007–08*

Teacher characteristics²	Mean/Frequency¹	
	Sanction All³	Sanction Title I only⁴
Assignment		
Elementary – tested	44.50	43.35
Elementary – not tested	7.02	9.44
Secondary - tested	26.53	24.81
Secondary - not tested	21.95	22.40
Experience		
Novice (less than or equal to 3 years)	17.21	15.06
More than 3 years	82.79	84.94
Sex		
Male	23.64	24.32
Female	76.36	75.68
Race		
All other races	18.82	6.36
White, non-Hispanic	81.18	93.64
Perception of leadership quality (factor score)	0.05	0.02
Years of experience ^{5, 6}	13.46	14.39
School characteristics		
AYP Status – Year 1 failure	25.06	21.00
Needs Improvement Status – Year 2 failure	12.40	10.14
Title I status	37.00	21.40
School level		
Elementary	72.23	69.50
Combined	5.15	7.64
Secondary	22.62	22.87
Enrollment (number of students) ⁷	589.95	478.61
Minority student population	42.78	25.32
Free- and reduced-price lunch population	47.60	38.80

¹Percentage unless otherwise stated.

²Minimum and maximum values equal full sample unless otherwise stated.

³Sanction: weighted statistics; unweighted sample sizes: teacher N=20,940; school N=4,250

⁴Sanction Title I Only: weighted statistics; unweighted sample sizes: teacher N=12,170; school N=2,330

⁵Sanction all: Minimum=1, Maximum=51

⁶Sanction Title I only: Minimum=1, Maximum=52

⁷Sanction Title I only: Minimum=0, Maximum=3,000

In contrast, numerous differences exist between the full sample and the non-sanction subsample, particularly at the school level. The non-sanction subsample contains slightly more teachers of elementary students in non-tested subjects, nearly 12 percent more White teachers, and teachers with slightly higher average years of experience. Compared to the analytic sample at the school level, the non-sanction sample contains smaller proportions of Title I schools and higher proportions of combined schools, as well as schools with lower average total, minority, and free- and reduced-price lunch enrollments. In sum, compared to the full sample, both subsamples differ by types of schools and by types of teachers, particularly by minority enrollment and the race of the teachers.

The two samples differ from one another along three teacher-level and all school-level characteristics. The non-sanction subsample has more teachers who teach elementary non-tested subjects, who are White, and who have, on average, an additional year of experience. But aside from a higher proportion of combined schools, all school-level characteristics are measurably lower for the non-sanction sample. Specifically, the non-sanction sample has smaller proportions of schools that do not make AYP, that are in Needs Improvement status, that are elementary schools, and that have lower average total, minority, and free- and reduced-price lunch enrollments. Simply put, on average, the non-sanction subsample has more high-performing schools with less diverse student populations taught by more White and more experienced teachers.

Multilevel model results. Table 14 presents the results of the multilevel model analysis for the full sample (model 3 from Table 10) and for each of the two subsamples. The coefficients for the sanctions subsample (model 4) are not statistically different from

the full model (model 3).^{38,39} Considering the general lack of descriptive differences between the full sample and the sanctions subsample, one would expect similarities between the two models. In fact, the only statistical difference between the full sample and the subsample for non-sanction states is the intercept. Essentially, the difference between the two intercepts indicates that compared to the full sample, the baseline value of teacher-perceived autonomy in the non-sanction sample is higher after controlling for all teacher and school characteristics in the model. Simply put, teachers in the non-sanction sample have higher average perceived autonomy than do those in the full sample.

³⁸ I made comparisons using dependent t-tests for statistical significance at the $p < .05$.

³⁹ Table 10 has models 1 through 3, with model 1 as the teacher-level model and model 3 as the HLM of the full sample.

Table 14

Multilevel Models for Scope and Degree of Teacher-Perceived Autonomy in the Classroom Domain by State-Level Sanction Policy^a

	Model 3	Model 4	Model 5
	Full sample	Sanction subsample	Non-sanction subsample
Intercept	0.40 ***	0.33 ***	0.53 ***
Teacher characteristics			
Elementary_Test subjects	-0.56 ***	-0.56 ***	-0.59 ***
Elementary_Non-tested subjects	0.18 ***	0.20 ***	0.09
Secondary_Test	-0.36 ***	-0.37 ***	-0.35 ***
Female	0.01	0.03	-0.04
Non-white	0.00	0.00	-0.05
Perception of leadership ^b	0.26 ***	0.27 ***	0.21 ***
Novice (le 3 years)	-0.14 ***	-0.15 ***	-0.11
Experience (years)	0.06 ***	0.05 ***	0.09 ***
School characteristics			
Fail AYP	-0.07 *	-0.05	-0.09 *
Needs Improvement	-0.01	0.00	-0.04
Schl level - Elementary	-0.31 ***	-0.30 ***	-0.34 ***
Schl level - Combined	0.12 **	0.06	0.17 **
Schoolwide Title 1	-0.07 *	-0.03	-0.14 **
Enrollment	-0.08 ***	-0.06 ***	-0.10 ***
% FRPL	0.00	-0.01	0.03
% Minority	-0.13 ***	-0.10 ***	-0.16 ***
Ave schl perception of leader ^b	0.19 ***	0.22 ***	0.13 ***
Interaction Terms			
Novice * level (Elem)	0.14 **	0.13 *	0.16 *
Novice * level (Combined)	-0.04	0.01	-0.07
Random Effects			
Intercept	0.18 ***	0.18 ***	0.15 ***
Perception of leadership	0.18 ***	0.19 ***	0.12 ***
Novice	0.32 ***	0.32 ***	0.28 ***

* $p \leq .01$; ** $p \leq .001$; *** $p \leq .0001$

^a Coefficients reported in effect sizes. All continuous measures were standardized across the entire sample ($M=0$; $SD=1$).

^b The perception of leadership measure is a factor score based on four categorical measures. The average school perception of leadership is the leadership factor score aggregated to the school level.

In contrast to the comparisons between the subsamples and the full sample, several differences exist between coefficients from the split models. Specifically, although estimates for teachers of elementary students in non-tested subjects and for novice teachers are not statistically significant in the non-sanction model, AYP status, combined schools, and schools with schoolwide Title I status are statistically significant. Thus, in the non-sanction subsample, on average teachers of elementary students in non-tested subjects do not differ from those who teach secondary students in non-tested subjects, and novice teachers do not differ from those who have average levels of experience. Further, on average teachers in schools that failed to make AYP perceive 9 percent of a standard deviation lower autonomy compared to those who work in schools that made AYP, and teachers in Title I schools perceive 14 percent of a standard deviation lower autonomy compared to those in non-Title I schools. Teachers in combined schools, on average perceive 17 percent of a standard deviation higher autonomy compared to those in secondary schools.

For those characteristics that maintained their statistical significance in both models, the magnitude of several point-estimates changed. In particular, teacher-level point-estimate differences include experience (increase from 0.05 SD to 0.09 SD) and perception of leadership (decrease from 0.27 SD to 0.21 SD). Differences in magnitude for school-level point-estimates include average total enrollment (absolute increase from -0.06 SD to -0.10 SD), average minority enrollment (absolute increase from -0.10 SD to -0.16 SD), and average school perception of leadership quality (decrease from 0.22 SD to 0.13 SD). Although the statistical significance of these differentials is not conclusive, the pattern of findings is noteworthy. That is, on average and compared to teachers who work

in schools that operate in sanction states, as experience increases teachers in schools in states that sanction only Title I schools perceive higher autonomy; and as average total and average minority enrollments increase, these teachers perceive lower autonomy. Further, teacher perception of school leaders at both the individual and school levels remains positive but has a reduced effect on perceived autonomy.

Summary of results

The findings detailed above encompass the teacher-level and school-level characteristics that affect teachers' perceptions of autonomy. At the teacher level, these characteristics include teaching assignment (tested versus non-tested), experience level (both novice status and years of experience), and perception of school leadership quality; at the school level, the characteristics that influence autonomy include the level taught (elementary, combined, and secondary), total and minority enrollment, perception of leadership quality, Title I status and AYP status.

Specific to the teacher level, across the full and subsample models, on average teachers who teach elementary students in tested subjects perceive nearly 60 percent of a standard deviation lower autonomy than do those who teach secondary students in non-tested subjects. The pattern contrasts sharply for elementary teachers of non-tested subjects. On average, teachers of elementary students in non-tested subjects perceive higher autonomy compared to secondary teachers in non-tested subjects; however, this relationship does not hold in the non-sanction subsample, where there is no difference for teachers in non-tested subjects by level. Across all three models, on average teachers who teach secondary students in tested subjects perceive about 36 percent of a standard deviation less autonomy compared to their peers who teach in non-tested subjects.

Meanwhile, novice teachers in elementary schools do not perceive autonomy differently than do those with between 3 and 14 years of experience, but novice and non-novice teachers alike perceive autonomy differently from non-novice teachers in secondary schools. The findings related to leadership, at the teacher and school levels, emphasize that school leaders have a moderate to large effect on teacher-perceived autonomy in the classroom. Teachers perceive higher autonomy in the classroom in schools where they perceive that the leaders are of high quality. Across all models, teacher-perceived autonomy is lower in schools with above average total school enrollment and percentage of minority enrollment.

Regarding AYP status and Title I status, key differences exist between the full and non-sanction models and the sanction model. Specifically, the full and non-sanction models indicate that teachers who work in schools that do not make AYP perceive lower autonomy compared to those who work in schools that did make AYP, while the sanction model indicates no difference on teacher-perceived autonomy based on AYP status. Similarly, compared to teachers who teach in lower-poverty schools, those who work in Title I schools perceive lower autonomy in both the full model and non-sanction models; whereas, teachers in Title I schools from the sanction model perceived no difference in autonomy from those who work in non-Title I schools.

The next and final chapter of this dissertation offers reflections on key findings for this study. Specifically, I discuss how these findings support or contradict prior research, what future researchers might learn from this study and advance in their own work, and how this study might inform policy.

Chapter 5: Discussion

This dissertation investigated hypotheses related to SBA reform and changes in teacher-perceived autonomy in the classroom. Specifically, I hypothesized that teacher-perceived autonomy had changed over time and that for teachers in 2007-08, certain teacher and school characteristics related to lower levels of autonomy. The research questions reflected these hypotheses. In this chapter, I discuss the overarching findings from the study and how the findings corroborate or contradict existing research. In general, I discuss the findings from the full sample and then findings from the subsample models. However, where findings hold across all models, I discuss the findings concurrently. I follow the findings discussion with implications for future research.

Discussion of the over-time hypothesis

The first question addresses the hypothesis that teacher-perceived autonomy has changed since the introduction of SBA reform policies in public schools and includes over-time and point-in-time perspectives. To address the over-time analysis, I use data from four waves of SASS (1993–94, 1999–2000, 2003–04, and 2007–08). I find that on average teachers perceive high levels of autonomy in the classroom. In fact, on a summative scale for each year that ranges from 6 to 24, most teachers fall near the 20 mark in each of the four data points. Although the averages offer a consistent pattern across time, they may be misleading for two reasons. First, one cannot directly compare the means across all data points because although the underlying psychometric properties are consistent across SASS administrations the scales are not equivalent between years. Second, just as NCLB brought to the fore the notion that average achievement values can mask differences between specific types of students (i.e., racial subgroups), average

values of teacher-perceived autonomy can mask variation across teacher types and school contexts.

To overcome the concerns regarding the average values, I employed CV and ICC statistics to determine whether the variation in teacher-perceived autonomy had changed over time. Indeed it had. Total variation in teacher-perceived autonomy increased from 1993 to 2008, as did variation due to school characteristics. In 1993-94, 12 percent of the variation in teacher-perceived autonomy was due to school context; whereas in 2007-08 nearly 21 percent was attributable to school context. Simply, over 15 years, school context has more influence on teachers' perceptions of their autonomy.

The finding that school context had a significantly increased role in 2007-08, compared to each of the three prior academic years may not be surprising. SBA reform is a top-down approach to standardize the education delivery system (Boyd, Lankford, Loeb, & Wyckoff, 2005; Manna, 2006b). By definition, the policymakers and organizations at the top would influence the policy recipients at the bottom. Further, the finding is consistent with Ingersoll's (2003) finding that schools influenced teacher autonomy through bureaucratic controls and mechanisms of teacher accountability. Indeed, if SBA implementation creates the desired outcomes, even in part, Dahl's power-influence constructs of scope, degree, and domain help us unpack how the organization (i.e., schools) increases influence over the individual (i.e., teacher) as policy implementation progresses. School-level curricular and accountability changes would widen the school-level scope of influence over the classroom domain and reduce the degree and scope of teacher influence over that domain, particularly if those changes

included areas historically left to teachers, such as student grading and evaluation or the teaching techniques employed.

Although I am able to offer empirical and conceptual evidence that associate the changes in teacher-perceived autonomy with SBA reform initiatives, the evidence for this association is circumstantial. The ICC models are not causal and do not control for other possible explanations of change. Thus, I cannot definitively state that SBA reforms led to the change in teacher-perceived autonomy during the time points between 1993 and 2008. Nonetheless, I can state the following: first, that between 2003-04 and 2007-08, the average perception of teacher autonomy fell slightly (based on mean values); second, that over time teachers increasingly varied from one another regarding perceptions of autonomy (based on the CV results); third, that school context increased in import from 1993-94 to 2007-08 related to teacher-perceived autonomy to indicate that forces other than teacher characteristics were involved in shaping teacher perceptions of their own autonomy (based on ICC results); and fourth, that the variability in teacher-perceived autonomy in Title I schools increased more so than it did compared to non-Title I schools (based on Title I ICC results). These four findings suggest a systemic change in public schools and one that may not be uniform in its implementation.

Discussion of the differential hypotheses

Specific to NCLB and 2007-08, I hypothesized that changes in teacher-perceived autonomy differed based on teacher and school characteristics, such as teacher experience, teacher assignment, school-level poverty, and school-level academic achievement. To complete these analyses, I used HLM to tease out specific school- and teacher-level characteristics associated with teacher-perceived autonomy, as well as to

isolate the influence of school-level effects on teacher characteristics. While I examined several teacher and school characteristics, I discuss the more salient topics in turn below.

Tested subjects and teacher-perceived autonomy. The findings in the full model, the sanction subsample model, and the non-sanction subsample model (models 3, 4 and 5, respectively) suggest that teacher-perceived autonomy differs substantially based on whether one teaches a tested subject. Again, examples of tested subjects are English and mathematics; examples of non-tested are arts and physical education. The difference made by teaching a tested subject is both the largest effect in the study and the most telling of what SBA reform has brought to schools. These models reveal large (from one-third to two-thirds of a standard deviation) and systemic differences between teacher-perceived autonomy based on teaching assignments. The finding that testing has an inverse relationship to teacher autonomy is congruent with prior scholarship that cites teachers' negative perceptions of the testing element of SBA (Abrams et al., 2003; Crocco & Costigan, 2007; Deniston and Gerrity, 2010; Luna and Turner, 2001; Mathison and Freeman, 2003); however, the magnitude of the finding is a contribution of this study.

In part, I interpret the large tested-subjects effect in this study as an indicator of curriculum effects, particularly in districts where schools leaders cooperate fully with external mandates (Hamilton, 2003; Valli, et al., 2008). That is, part of the large effect size for tested subjects may reflect the overall curricular changes that result from the contents and demands of the standardized test regimen. To illustrate, a district mandates a standardized curriculum in all tested subjects. In this district an arts teacher (non-tested subject) may not face the same level of SBA reform enforcement or scrutiny or pressure

as a teacher in a tested subject such as science or math. While an arts teacher may perceive general changes at the school, a science teacher—whose students are tested—may experience direct effects of SBA reform. Rather than offering teachers suggested instructional lessons in science, the district SBA-inspired curriculum may prescribe curricular content and instructional methods for the science teacher. In this district, and based on the definition of scope of power in this study,⁴⁰ the tested-subject effect would reflect narrowed scope of power over the classroom domain for teachers of tested subjects in the areas of textbook selection; teaching technique selection; and content, topics, and skills selection at the very least. For the science teacher, a diminished scope of power over these areas would decrease his or her perception of autonomy in the classroom domain in ways that might not affect the arts teacher.

I contend that the tested-subject effect also demonstrates a circumstance of inequity in the SBA reforms, as adopted and applied in NCLB. Given the stark differential effect on perceived autonomy between teachers in non-tested subjects and those in tested subjects, the tested-subject effect exposes a disproportionate application of accountability such that teachers who opt to teach core subjects (as defined by NCLB) perceive declines in classroom autonomy where teachers who do not teach core subjects do not perceive a decline in classroom autonomy. The condition where some teachers perceive lower autonomy based on assignment seems not only unsustainable from a labor market perspective, but also counterproductive for the standards movement. Akin to earlier federal education policies, NCLB does target specific student populations (e.g., high-poverty). However, it seems that the perception of lower autonomy in the classroom

⁴⁰ From the introduction chapter, scope includes: textbook selection; teaching technique selection; content, topics, and skills selection; student grading and evaluation; student discipline; and homework determination

by specific types of teachers based on assignment should not be part of the targeted effort encapsulated in NCLB. If SBA reforms aspire to ensure consistent content and performance standards for students, difficulty may arise if the delivery of that standardized education disproportionately affects teachers based on assignment.

Experience levels and teacher-perceived autonomy. A second teacher-level finding regarding the differential hypotheses relates to novice status and experience in general. Similar to findings from studies reviewed in chapter 2, findings for novice teachers in the full and sanction models in this study indicate that novice teachers perceive lower autonomy than their non-novice colleagues based not only on novice teachers' inexperience, but also on the level of the school in which they teach. Novice teachers in secondary schools perceive lower autonomy compared to their non-novice peers, while all elementary school teachers perceive lower autonomy compared to non-novice secondary school teachers. The experience effect, which holds across all models, complements the novice findings because for teachers with average experience (13.66 years), as their experience increases so does their perceived autonomy. The findings regarding levels of experience and grade level are congruent with prior studies in which newer teachers and elementary school teachers perceive lower autonomy than their colleagues with more experience or their peers who teach in secondary schools (Allensworth, E., Ponisciak, S., & Mazzeo, C; Perrow; and Pearson, 1995). I suggest that the findings related to novice status and average experience may be more a function of overall levels of experience than of SBA reform, particularly since the larger effect of inexperience is at the secondary level where SBA reform policies tend to be less intensive than they are in the early grades.

In contrast, the finding that elementary school teachers perceive lower autonomy compared to those who teach in secondary schools aligns with the testing effect because NCLB requires annualized standardized tests in third through eighth grade but makes fewer testing demands in higher grades. Further, the finding at the elementary level is consistent across the three models. That is, on average teachers in elementary schools perceive lower autonomy than those in secondary schools. The effect of the elementary school level on teachers substantiates the findings from the ICC analyses that school context plays an important role in teacher-perceived autonomy. The power-influence framework helps explain the elementary school finding. If schools widen their scope and augment their degree of power over the classroom domain to implement NCLB while teachers concomitantly experience a decrease in their perceived autonomy, then one might explain the perception of decreased autonomy as a diminished scope and diminished degree of power over the classroom domain due to competitive claims on school goals (Leiter, 1981).

School leadership and teacher-perceived autonomy. A third salient teacher-level finding is the large effect for teacher perception of school leadership. The magnitude of the leadership effect is noteworthy. Consistently, the models indicate that as teachers' favorable perception of their leader increases, so does their perceived autonomy. This finding aligns with what one would expect given prior studies. Specifically, prior research has documented that principals can protect teachers from external influences that may alter teachers' scope of influence. Conversely, principals can use external influences (e.g., SBA reform) to implement school change that teachers

deem intrusive because of realized or potential narrowed teacher scope over the classroom domain (Allensworth et al., 2009; Hanson, 2006; Leiter, 1981).

The school-level leadership effect bolsters the teacher-level leadership effect because even after controlling for individual teacher perceptions of leadership, in each model the collective school-average perception of leadership influences teacher-perceived autonomy. As average school-level favorable perceptions of the school leader increase, so does perceived autonomy. During the NCLB implementation process, it is reasonable to suggest that principals have become even more central as arbiters of external policy. The positive association of both individual and collective school leadership to teacher-perceived autonomy can play out in at least two scenarios: one, in which teachers support the changes that principals enact; two, in which teachers value their principals who, for good or for ill, avoid and buffer external intervention where possible.

Although the leadership effect remains positive, differences between the sanction and non-sanction models may reflect back to the tested-subject effect discussed above. It is noteworthy that in the non-sanction model the leadership effect at both the teacher and school level is lower than it is in the sanction model, while the tested-subject effect remains at nearly 60 percent of a standard deviation.⁴¹ Effectively, school leadership has less influence on teacher-perceived autonomy in non-sanction states but the tested-subject effect holds its large and negative relationship with teacher-perceived autonomy for both subsamples. It is possible that teachers who work in schools in non-sanction states either do not feel as protected from external demands, such as adoption of district curriculum; or that teachers in these schools perceive a disorganized or ineffectual response from

⁴¹ I made comparisons using independent t-tests for statistical significance at the $p < .05$.

their school leaders. Thus, while still important to teachers, leadership is not as influential to teacher-perceived autonomy in these schools as it is in the sanction sample.

This study does not tease out the motivation behind the relationship between leadership and teacher-perceived autonomy. However, it documents the extension of the leadership effect using the nationally representative 2007-08 data to corroborate existing evidence that suggests that principals matter a great deal to teacher autonomy. In sum, teachers perceive higher levels of autonomy in schools where they perceive high-quality leaders.

Title I status and teacher-perceived autonomy. Title I findings provide compelling school-level findings regarding the notion that teachers in different types of schools experience autonomy differently, particularly when one examines the ICC and HLM findings in tandem. Figure 7 shows the ICC for Title I schools as 23 percent in 2007-08, which was higher than prior years and higher than non-Title I schools. Model 3 (the full sample) in Table 14 lists a negative effect (-.07SD) for teacher-perceived autonomy for teachers who work in Title I schools. The effect is even stronger for those who work in Title I schools in states that sanction only Title I schools (-.14SD). Taken together, these findings indicate that the increase in variance between Title I schools in 2007-08 corresponds with a lower average teacher perception of autonomy in Title I schools that operate in states that sanction only Title I schools. Simply, in states that do not apply accountability measures (i.e., sanctions) more broadly across all schools, teachers in Title I schools perceive lower autonomy.

The effect of Title I status is important from a policy standpoint because, like the testing results, it highlights how the implementation of SBA reforms such as NCLB may

result in a system of accountability that targets certain teachers and schools to the exclusion of others. While accountability and standardization are goals of SBA in general and NCLB specifically, the privileging or targeting of some subjects or school settings over others may have real effects on teacher-perceived autonomy and, in turn, on teacher job satisfaction and turnover. Further evidence that an inequitable application of accountability may be at play is evidenced by the lack of difference in teacher-perceived autonomy between teachers in Title I schools and non-Title schools in those states that sanction all schools regardless of Title I status. Effectively, these states have balanced the accountability equation. Policymakers may want to consider the findings related to Title I (and AYP discussed in detail below) between the sanction and non-sanction models as they work toward ESEA reauthorization.

AYP Status and teacher-perceived autonomy. Both parts of the second research question concern AYP status, the direct measure of SBA reform in NCLB. At the start of this study, I anticipated a moderate to large AYP effect on teacher-perceived autonomy. Further, I hypothesized that the AYP effect would hold, if not increase, for schools in need of improvement (i.e., in the sanction phase of NCLB). To the contrary, the AYP effect is small and inconsistent and the Needs Improvement effect is nonexistent. Although on average, AYP has an inverse relationship with teacher-perceived autonomy for the full sample, given the results from the subsample analyses it appears that the inverse relationship is in part due to the variance between teachers in sanction states and those in non-sanction states. Specifically, teachers in sanction states who work in schools that fail to make AYP perceive autonomy no differently from those who work in schools that make AYP. Alternatively, teachers in non-sanction states who

work in schools that failed AYP perceive slightly less autonomy than those who work in schools that met AYP.

Both the pattern of the AYP effect and the relatively small magnitude (less than 10 percent of a standard deviation in both models) of the effect are surprising. The sanction subsample contains all states that have decided to apply NCLB sanctions to schools that fail AYP regardless of Title I status. Hence, the null finding regarding Title I status in this subsample makes sense. However, I expected teachers in these states to perceive lower autonomy from AYP given that the threat of sanctions applies to all schools. Compared to prior renditions of ESEA, NCLB brings more intense federal focus to accountability through sanctions. The power-influence constructs might indicate that federal actors have widened their scope of power over schools by using the sanction lever. Again, the model did register an effect (although not causal) for AYP status and teacher-perceived autonomy in the full sample, but given the state policy regarding sanctions, the findings in the two subsample models are counter-intuitive. Before I explore the possible explanations for the AYP effects (or lack thereof), I will discuss the Needs Improvement results because it is reasonable that the explanations for these two sets of AYP findings relate to one another.

Findings regarding to the Needs Improvement AYP status were surprising based on my hypothesis that teachers who worked in schools that failed to make AYP would perceive lower autonomy compared to those who worked in schools that did not fail to make AYP. After controlling for teacher- and school-level characteristics, as well as AYP status in the first year, teachers who worked in schools that failed to meet AYP for two or more consecutive years did not perceive autonomy any differently from those who

worked in schools that met AYP. Simply, AYP made a small difference to teacher perceptions in a specific setting but only in the first year of failure. I offer three suggestive scenarios grounded in the reviewed literature and theoretical framework of this study to explain the AYP and Needs Improvement findings.

First, I suggest the low effect, or lack thereof, for AYP and Needs Improvement might be a signal of capacity. Elmore (2002) explains that the result of historic state and local control over education is that respective resources, activities, and efforts did not build a cohesive institutional capacity to “monitor and improve teaching and learning,” (p. 35). In other words, states, districts, and schools separately built the capacity needed to either deliver education, assess education, or some combination thereof, but the local level capacity building may have impeded a coherent system wide ability to improve education. Thus, no capacity that spans state, local, and school levels exists to translate assessment into improvement for teachers or students. For example, states focused on instructional delivery through monitoring of teacher education programs; districts focused on professional development programs, staffing, and curricula. School staff worked to master instruction, whether through adoption and delivery of curricula or through strategies that would lead to strong tests scores. In states that opted to sanction any school that failed AYP, it is reasonable to consider that where they had high levels of failure (such as Kentucky and California), the SEA was not able to muster enough resources to follow through on the enforcement of sanctions in all schools that failed AYP.

System-wide capacity issues may influence teacher-perceived autonomy such that once the threat of sanctions becomes reality and schools face continued years of sanctions

with minimal disruption to the workplace, teachers do not perceive diminished autonomy in the classroom because their scope and degree of autonomy remains as it had been. The realization at the local level that the state level lacked capacity quite possibly squelches the sting of the failed AYP status. Similarly, if teachers work to avoid AYP failure but despite their best efforts the school did not make AYP, teachers may perceive that the district, the state, or the combination thereof did not have the organizational capacity to raise scores. Teachers may externalize the failure. For example, teachers may perceive school failure as the district lacking capacity to offer training and supports that teachers might need; and that the SEA did not have the financial resources, instructional expertise, or qualified personnel to target at these schools. Thus, teachers would not internalize continued failure because the fault may lie with issues beyond their purview. In this case, teachers would conduct their classes with the same levels of perceived autonomy as they had prior to the Needs Improvement phase of AYP.

A second scenario involves the position that states take regarding NCLB implementation. Since the first year of NCLB, states have sought waivers to opt out of or transform aspects of the Act related to definitions and assessment of AYP status.⁴² I hypothesize that regardless of the outcome of the waiver request, the availability of said process evokes the Cuban (1990) perspective of education reform in which state, district, and school staff recognize that the education reforms, including NCLB, are not resolute because of the decoupled nature of the education system. Further, like many reforms that came before, the more contested notions of NCLB (i.e., AYP sanctions) will either

⁴² While this study alluded to recent success for state related to waivers as part of the historical context of SBA reform the findings from 2007–08 data would not reflect these 2011 developments.

change to fit prior practice or will dissolve altogether due to the lack of state-level compliance that federal actors need to legitimize and implement the Act.

The final scenario, and perhaps the most likely, is that the sanctions most commonly implemented for AYP failure do not influence teacher perceptions of autonomy. Detailed in Table 1, the first year sanction is school choice for students in the failed school; the second year sanction adds supplemental education services.⁴³ The school transfer sanction is a district- and school-level sanction that would not necessarily register with an autonomy construct confined to classroom activities like textbook selection and homework assignment. Similarly, the supplemental services sanction is at the school level in that school leaders must find providers to deliver the services which take place outside of classroom hours. Thus, quite possibly teachers do not perceive any change to their autonomy in the classroom domain due specifically to NCLB sanctions that schools enact outside of that domain.

Concluding remarks.

Taken together, the findings in this study indicate that SBA reform does indeed influence teacher-perceived autonomy. The theory of action of SBA is that systemic reform which aligns standards, content, and assessment will lead to measurable improvement in student achievement. Testing is an integral part of SBA reform, be it low- or high-stakes, federal or state sponsored. Considering SBA's focus on standardization, curricular alignment, and accountability, the findings from this study, such as those for teaching in tested subjects and Title I schools, indicate that SBA reforms, and NCLB specifically, have influences on public education that include

⁴³ While the sanctions extend to staff termination and school reconstitution, I think it imprudent to imagine that by 2007–08 many teachers worked in schools that were much beyond their third year of consecutive failure.

changes to teacher-perceived autonomy in the classroom. However, the study found a differential effect of SBA's influence on teacher-perceived autonomy that may be an unintended consequence due to the structure of NCLB.

Specifically, NCLB targets low-performing schools and has a particular emphasis on Title I schools. Thus, in these schools lower perceptions of teacher autonomy would be not only expected but also necessary according to the policy. However, the findings here suggest that NCLB influences teacher-perceived autonomy in *all* schools based on grade and subject matter taught. For example, the large tested-subject effect contrasted with the relatively small and null findings related to AYP status and Needs Improvement status may indicate that the more immediate curricular and testing requirements result in lower perceptions of autonomy in the classroom, rather than the possibility of NCLB sanctions. However, NCLB's disparate approach to subject matter and teaching assignments, in addition to Title I school context, offer a compelling example of how SBA in general and NCLB in particular continue the tension between autonomy and accountability in public schools.

Implications for Future Research

Given the findings of and limitations to the current study, several opportunities for future research exist. In this section I offer four ideas for future research.

First, considering the testing effect in this study, future researchers may want to use multiple waves of SASS data, including the soon-to-be released 2011-12 data, to create multilevel models that include the testing indicators to investigate whether the 2007–08 finding is the culmination of an over-time trend, is specific to post NCLB waves (2003–04 and 2007–08), or is unique to this study. Based on the ICC growth over time, it is likely that an over-time HLM analysis would produce intriguing results. Further, researchers should consider combining the SASS data with the retention data available through the Teacher Follow-Up Survey (TFS) in SASS. A study that married the autonomy construct with the mobility data available in TFS may be able to determine what effect, if any, a change in teacher-perceived autonomy has on retention.

The second idea responds to the single method approach in this study. That is, using SASS or some other reliable teacher data as a baseline to identify characteristics that influence teacher autonomy, researchers should consider turning to qualitative methods to capture teacher voices and impressions. Of particular interest might be the definition that teachers would offer for autonomy. Do teachers interpret autonomy in broad or narrow terms, and how do individual teacher definitions vary across teachers' characteristics and school contexts in which they work? Do teachers put different value on autonomy related to *what* is taught or to *how* to teach? These key questions that capture tension between the SBA theory of action and the teacher autonomy would concomitantly permit a more fulsome and precise measure of teacher-perceived

autonomy that separates the *what* from the *how* and that would be a valuable construct for future research on the intersection of school reform and teaching.

The findings related to principals offer a third avenue of research. A mixed-methods study of principals could investigate principal beliefs, practices, and attitudes about their roles as instructional leaders, as staff managers, and as community leaders. This type of study would provide deep insight into how and why principals are so important to teacher-perceived autonomy. Again, using SASS or similar data one could ground the study and extend it through case study methods that permit in-depth analysis and understanding of principals. Given their role as intermediaries and negotiators between federal, state, and district policies and teacher practices, a study that examines whether principals perceive any differential effects in their construct of autonomy over schools based on SBA reform would be a robust extension of the current study.

Fourth, I suggest a study that would seek to identify the conditions under which the public school system should foster teacher autonomy and when the system should restrict teacher autonomy. The data in this study would not support a qualitative investigation of teacher autonomy (e.g., Should teachers have power over the classroom? And if so, how much), but researchers can collect data from actors across levels of the system to ascertain information across the spectrum of teacher autonomy. For example, researchers might ask teachers under what circumstances teachers respond to requirements to follow curricula guides, disciplinary protocol, or district-wide textbook adoptions. Researchers might ask teachers under what circumstances principals should dictate classroom activity. Similarly, researchers might ask principals and district officials under what circumstances teachers should and should not make decisions

regarding instructional practices and delivery; and under what circumstances local and state level policies should dictate classroom practices.

Appendix

Table A. 1

Factor Loadings

	1993 (N=44040)	1999 (N=38740)	2003 (N=40280)	2007 (N=33110)
Factor variables				
select text	0.5640	0.5022	0.4990	0.5295
select content to be taught	0.6619	0.6093	0.6003	0.6001
teaching technique	0.7495	0.7609	0.7428	0.6954
evaluate/grade students	0.6760	0.7236	0.7129	0.6713
decide discipline	0.5211	0.4990	0.5114	0.4193
assign homework	0.5753	0.6047	0.5945	0.5442
Order of factor loadings				
1 Teach	1 Teach	1 Teach	1 Teach	1 Teach
2 Evaluate	2 Evaluate	2 Evaluate	2 Evaluate	2 Evaluate
3 Content	3 Content	3 Content	3 Content	3 Content
4 Homework	4 Homework	4 Homework	4 Homework	4 Homework
5 Text	5 Text	6 Discipline	5 Text	5 Text
6 Discipline	6 Discipline	5 Text	6 Discipline	6 Discipline

Table A. 2

Eigenvalues and Percentage of Variance Explained

	1993 (N=44040)	1999 (N=38740)	2003 (N=40280)	2007 (N=33110)
Eigenvalues	3.4939	3.3051	3.2889	2.6457
Percent of variance	0.5823	0.5509	0.5482	0.4409

Table A. 3

Results from Scale and Tertile Factor Outcome Variables

	Scale	Tertile Factor ^b	
	Estimate	High Estimate	Low Estimate
Intercept	0.42 ***	-0.14 ***	0.80 ***
Teacher characteristics			
Elementary_ Tested subjects	-0.56 ***	-2.05 ***	-0.94 ***
Elementary_ Non-tested subjects	0.18 ***	0.12	-0.06
Secondary_ Tested	-0.37 ***	-1.20 ***	-0.35 ***
Novice (le 3 years)	-0.05 **	-0.16	0.07
Experience (years)	0.06 ***	0.19 ***	0.02
Female	0.01	-0.07	-0.03
Non-white	-0.01	-0.19	0.12
Perception of leadership	0.26 ***	0.68 ***	0.42 ***
School characteristics			
Fail AYP	-0.07 **	-0.19	-0.04
Schl level - Elementary	-0.32 ***	-0.83 ***	-0.45 ***
Schl level - Combined	-0.09 *	0.24	0.06
Title 1 status	-0.05 *	-0.11	-0.14
Enrollment	-0.07 ***	-0.25 ***	-0.07 *
% FRPL	0.00	0.01	0.03
% Minority	-0.13 ***	-0.36 ***	-0.20 ***
Ave schl perception of leader	-0.07 ***	-0.22 *	-0.08
Random Effects			
Intercept	0.16 ***	0.76 ***	0.61 ***

* $p \leq .01$; ** $p \leq .001$; *** $p \leq .0001$

^a Coefficients for the two factor models are odds ratios. The HIGH model coefficients can be interpreted as the cumulative odds of belonging to the high autonomy group (1 =high) vs. belonging to moderate or low autonomy groups (moderate=2, low=3). The MODERATE model coefficients can be interpreted as the cumulative odds of belonging to the low autonomy group. Positive coefficients increase the probability of being in the modeled group; negative coefficients decrease the odds. For example, the odds of being in the HIGH autonomy group for teachers who teach elementary students in tested subjects is two times *less* likely than those who teach secondary students in non-tested subjects. Coefficients for the zscale model can be interpreted as effect sizes. All continuous measures were standardized ($M=0$; $SD=1$).

^b Tertile was created through factor analysis. The factor was categorized as 25% high, 25% low and 50% moderate.

Table A. 4

Perception of Leadership Factor Pattern, Eigenvalue, and Variance Explained

Factor variables	Pattern
support	0.7736
rules	0.7761
communicates	0.7566
recognition	0.6940
Statistics	
Eigenvalue	3.3628
Variance explained	0.8407

Table A. 5

Sanction Policies by State

	Sanctions for all schools that fail AYP	Sanctions for only Title I schools that fail AYP
Alabama	x	
Alaska		x
Arizona	x	
Arkansas	x	
California	x	
Colorado	x	
Connecticut		x
Delaware	x	
District of Columbia		x
Florida	x	
Georgia	x	
Hawaii	x	
Idaho	x	
Illinois	x	
Indiana	x	
Iowa		x
Kansas		x
Kentucky	x	
Louisiana	x	
Maine		x
Maryland	x	
Massachusetts	x	
Michigan	x	
Minnesota		x
Mississippi	x	
Missouri		x
Montana		x
Nebraska		x
Nevada		x
New Hampshire		x
New Jersey		x
New Mexico	x	
New York	x	
North Carolina	x	
North Dakota		x

Table A. 5: Sanction policies by state (cont.)

Ohio	x	
Oklahoma	x	
Oregon		x
Pennsylvania	x	
Rhode Island	x	
South Carolina	x	
South Dakota		x
Tennessee	x	
Texas	x	
Utah		x
Vermont	x	
Virginia	x	
Washington		x
West Virginia	x	
Wisconsin		x
Wyoming	x	
U.S. Total	32	19

Table A.6

Intraclass Correlation Coefficients (95% Confidence Intervals), 1993–2007

	1993 (n = 44040)	1999 (n = 38740)	2003 (n = 40280)	2007 (n = 33110)
Total sample	0.1209 (0.1206, 0.1211)	0.1393 (0.1389, 0.1396)	0.1712 (0.1709, 0.1716)	0.2078 (0.2074, 0.2083)
Title I	0.1244 (0.1239, 0.1247)	0.1482 (0.1476, 0.1487)	0.1603 (0.1598, 0.1607)	0.2303 (0.2297, 0.2310)
Non-Title I	0.1163 (0.1158, 0.1166)	0.1272 (0.1267, 0.1276)	0.1701 (0.1694, 0.1705)	0.1799 (0.1794, 0.1804)

Table A.7

CV and ICC Calculations

	1993 (n=44040)	1999 (n=38740)	2003 (n=40280)	2007 (n=33110)
<i>Autonomy scale¹</i>				
Mean (wgt)	20.251	19.610	20.314	19.740
Std Dev (wgt)	21.429	25.373	26.815	30.205
CV	1.0582	1.2939	1.3200	1.5301
Intercept (Tau)	0.934	1.187	1.4952	1.9466
Residual (Sigma sq)	6.793	7.337	7.2397	7.5030
ICC	0.1209	0.1393	0.1712	0.2060

Table A. 8

CV and ICC by Title I Status, 1993–2007

	1993 (n=44040)	1999 (n=38740)	2003 (n=40280)	2007 (n=33110)
<i>Autonomy scale¹ - Title I Only</i>				
Intercept (Tau)	0.986	1.347	1.5972	2.3872
Residual (Sigma sq)	6.937	7.743	7.7939	8.1382
ICC	0.1244	0.1482	0.1701	0.2268
<i>Autonomy scale - Non-Title I Only</i>				
Intercept (Tau)	0.874	1.027	1.3169	1.5480
Residual (Sigma sq)	6.645	7.044	6.9008	7.0994
ICC	0.1163	0.1272	0.1603	0.1799

¹Unstandardized

References

- Abrams, L.M., Pedulla, J., & Madaus, G. (2003). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory into Practice*, 42(1), 18-29.
- Airasian, P. (1988). Symbolic validation: The case of state-mandated, high-stakes testing. *Educational Evaluation and Policy Analysis*, 10(4), 301-313.
- Allensworth, E., Ponisciak, S., & Mazzeo, C. (2009). The schools teachers leave: Teacher mobility in Chicago Public Schools (pp. 1-52). Chicago: Consortium on Chicago School Research, University of Chicago.
- Allington, R.L. (2006). Reading lessons and federal policy making: An overview and introduction to the special issue. *The Elementary School Journal*, 107(1), 3-15.
- Amrein, A.L., & Berliner, D.C. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, 10(18), 1-74.
- Archbald, D., & Porter, A. (1994). Curriculum control and teachers' perceptions of autonomy and satisfaction. *Educational Evaluation and Policy Analysis*, 16(1), 21-39.
- Barker, N. (2005). *A practical introduction to the bootstrap using the SAS system*. Oxford Pharmacueticals Sciences. Wallingford, UK.
- Bathon, J., & Spradlin, T. (2007). Outcomes of the school choice and supplemental educational services provisions of NCLB. *Eduation Policy Brief*, 5(8).
- Beteille, T., Kalogrides, D., & Loeb, S. (2009). Effective schools: Managing the recruitment, development, and retention of high-quality teachers. CALDER Working Paper 37. Washington, DC: The Urban Institute.
- Birman, B., Le Floch, K.C., Klekotka, A., Ludwig, M., Taylor, J., Walters, K. (2007). State and local implementation of the No Child Left Behind Act: Volume II -- Teacher quality under NCLB: Interim report. In R. Corporation (Ed.). Santa Monica, CA.
- Bogler, R., & Somech, A. (2004). Influence of teacher empowerment on teachers' organizational commitment, professional commitment and organizational citizenship behavior in schools. *Teaching and Teacher Education*, 20, 277-289.
- Booher-Jennings, J. (2005). Below the bubble: "Educational Triage" and the Texas accountability system. *American Educational Research Journal*, 42(2), 231-268.

- Boote, D. (2006). Teachers' professional discretion and the curricula. *Teachers and teaching: theory and practice*, 12(4), 461-478.
- Borkowski, J., & Sneed, M. (2006). Will NCLB Improve or Harm Public Education? *Harvard Educational Review*, 76(4), 503-525.
- Boyd, D., Grossman, P., Ing, M., Lankford, H., Loeb, S., & Wyckoff, J. (2011). The influence of school administrators on teacher retention decisions. *American Educational Research Journal*, 48(2), 303-333.
- Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). Explaining the short careers of high-achieving teachers in schools with low-performing students. *American Economic Review*, 95(2), 166-171.
- Brown, C., & Cooper, B. (2011). Education politics and equity policy since 1950: From rights to accountability. In D. Mitchell, R. L. Crowson & D. Shipps (Eds.), *Shaping education policy* (pp. 143-163). New York, NY: Routledge.
- 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.
- Burbules, N.C., & Densmore, K. (1991a). The limits of making teaching a profession. *Education Policy*, 5(1), 44-63.
- Burbules, N.C., & Densmore, K. (1991b). The persistence of professionalism: Breaking up is hard to do. *Educational Policy*, 5(2), 150-157.
- Campbell, E. (2006). Editorial: Curricular and professional authority in schools. *Ontario Institute for Studies in Education of the University of Toronto*, 36(2), 113-118.
- Campbell, R., Cunningham, L.L., Nystrand, R.O., & Usdan, M.D. (1985). *The organization and control of American schools* (5th ed.). Columbus, OH: Merrill Publishing Company.
- Campbell, R.F., Cunningham, L.L., Nystrand, R.O., & Usdan, M.D. (1990). *The organization and control of American schools* (6th ed.). New York, NY: Macmillan Publishing Company.
- Campbell, R.F., & Mazzoni, T.L. (1976). Conceptual framework and research methodology. *State policy making for the public schools* (pp. 3-13). Berkeley: McCutchan.
- Clotfelter, C.T., Ladd, H.F., & Vigdor, J.L. (2006). High poverty schools and the distribution of teachers and principals: Sanford Working Paper Series San06-08.

- Clune, W.H. (1993). The best path to systemic educational policy: Standard/centralized or differentiated/decentralized. *Educational Evaluation and Policy Analysis*, 15(3), 233-254.
- Cohen-Vogel, L. (2005). Federal role in teacher quality: "Redefinition" or policy alignment? *Educational Policy*, 19(1), 18-43.
- Cohen, D. (1996). Rewarding teachers for student performance. In S. Fuhrman & J. O'Day (Eds.), *Rewards and reform* (pp. 60-112). San Francisco: Jossey Bass.
- Cohen, D., Moffitt, S., & Goldin, S. (2007). Policy and practice: The dilemma. *American Journal of Education*, 113(4), 515-548.
- Conley, S. (1991). Review of research on teacher participation in school decision making. *Review of Research in Education*, 17, 225-266.
- Crocco, M.S., & Costigan, A.T. (2007). The narrowing of curriculum and pedagogy in the age of accountability: Urban educators speak out. *Urban Education*, 43(6), 512-535.
- Cuban, L. (1990). Reforming again and again. *Educational Researcher*, 19(1), 3-13.
- Dahl, R. (1984). *Modern political analysis* (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Darling-Hammond, L. (2004a). From 'Separate but Equal' to 'No Child Left Behind': The collision of new standards and old inequities. In D. Meier & G. Wood (Eds.), *Many children left behind: How the No Child Left Behind Act is damaging our children and our schools* (pp. 3-32). Boston: Beacon Press.
- Darling-Hammond, L. (2004b). Standards, accountability, and school reform. *Teachers College Record*, 106(6), 1047-1085.
- Darling-Hammond, L., Berry, B., & Thoreson, A. (2001). Does teacher certification matter? Evaluating the evidence. *Educational Evaluation and Policy Analysis*, 23(1), 57-77.
- Darling-Hammond, L., & Wise, A.E. (1985). Beyond standardization: State standards and school improvement. *The Elementary School Journal*, 85(3), 315-336.
- Davis, S., Darling-Hammond, L., LaPointe, M., & Meyerson, D. (2005). *School leadership study: Developing successful principals* (Review of Research). Stanford, CA: Stanford University, Stanford Educational Leadership Institute.
- DeBray, E. (2005). Partisanship and Ideology in the ESEA Reauthorization in the 106th and 107th Congresses: Foundations for the new political landscape of federal education policy. *Review of Research in Education*, 29, 29-50.

- DeBray, E. (2006). *Politics, ideology, and education: Federal policy during the Clinton and Bush administrations*. New York: Teacher's College.
- DeBray, E., McDermott, K., & Wohlstetter, P. (2005). Introduction to the special issue on federalism reconsidered: The case of the No Child Left Behind Act. *Peabody Journal of Education*, 80(2), 1-18.
- Deci, E., & Ryan, R. (1987). The support of autonomy and control of behavior. *Journal of Personality and Social Psychology*, 53(6), 1024-1037.
- Deniston, R.D., & Gerrity, K.W. (2010). Elementary school teachers' perceptions of No Child Left Behind and its effect on morale. *scholarlypartnershipsedu*, 5(2), 26-34.
- Diamond, J., & Spillane, J. (2004). High-stakes accountability in urban elementary schools: Challenging or reproducing inequality? *Teachers College Record*, 106(6), 1145-1176.
- Dorn, S. (1998). The Political Legacy of School Accountability Systems. *Education Policy Analysis Archives*, 6(1).
- Downie, R.S. (1990). Professionas and professionalism. *Journal of Philosophy of Education*, 24(2), 147-159.
- Edelman, L.B. (1992). Legal ambiguity and symbolic structures: Organizational mediation of civil rights law. *American Journal of Sociology*, 97(6), 1531-1576.
- Eisner, E. (1995). Standards for American Schools. *Phi Delta Kappan*, 76(10), 758-764.
- Elmore, R. (2002). Unwarranted intrusion. *Education Next*, Spring(2002), 31-35.
- Elmore, R., & McLaughlin, M. (1982). Strategic choice in federal education policy: The compliance-assistance tradeoff. In A. Lieberman & M. McLaughlin (Eds.), *Policymaking in education, 81st yearbook of the National Society for the Study of Education* (pp. 159-194). Chicago: University of Chicago Press.
- Elmore, R.F. (1987). Reform and the culture of authority in schools. *Educational Administrative Quarterly*, 23(4), 60-77.
- Elmore, R.F., Abelman, C., & Fuhrman, S. (1996). The new accountability in state education policy. In H. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp. 65-98). Washington D.C.: Brookings Institution.
- Feng, L., Figlio, D., & Sass, T. (2010). School Accountability and Teacher Mobility. CALDER Working Paper 47. Washington, DC: The Urban Institute.

- Finnigan, K., & 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. (2001). *The governance of teaching and standards-based reform from the 1970s to the new millenium*. Paper presented at the The Annual Meeting of the Center on Research on Educational Opportunity: Conference on Stability and Change in Education, Notre Dame, IN.
<http://www.cepa.gse.rutgers.edu/Shortcut%20to%20Governabilitybib.lnk.pdf>.
- Firestone, W., & Bader, B. (1991). Professionalism or bureaucracy? Redesigning teaching. *Educational Evaluation and Policy Analysis*, 13(1), 67-86.
- Firestone, W., Fuhrman, S., & Kirst, M.J. (1991). State educational reform since 1983: Appraisal and the future. *Educational Policy*, 5(3), 233-250.
- Firestone, W., 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., & Pennell, J. (1993). Teacher commitment, working conditions, and differential incentive policies. *Review of Educational Research*, 63(4), 489-525.
- Forte, E., & Erpenbach, W.J. (2006). 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 *Statewide educational accountability under the No Child Left Behind Act*. Washington, D.C.: CCSSO
- Fuhrman, S. (1999). The new accountability. CPRE Policy Brief (pp. 3-11). Philadelphia: Consortium for Policy Research in Education
- Fuhrman, S., & Elmore, R.F. (1990). Understanding local control in the wake of state education reform. *Educational Evaluation and Policy Analysis*, 12(1), 82-96.
- Gamse, B., Jacob, R.T., Horst, M., Boulay, B., & Unlu, F. (2008). Reading first impact study final report (NCEE 2009-4038). In NCEE (Ed.). Washington DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Services, U.S. Department of Education.
- GAO. (2006). *No Child Left Behind Act: Education actions needed to improve local implementation and state evaluation of supplemental educational services*. 06-758 Washington, D.C.
- Gawlik, M. (2007). Beyond the charter schoolhouse door: Teacher-perceived autonomy. *Education and Urban Society*, 39(4), 524-553.

- Geary, L.S. (1992). *The policymaking process resulting in fiscal policy for special education in the state of Utah*. Educational Administration. Unpublished manuscript. University of Utah. Salt Lake City.
- Glazer, J. (2008). Educational professionalism: An inside-out view. *American Journal of Education*, 114(February), 169-189.
- Goertz, M.E. (2001). Redefining Government Roles in an Era of Standards-Based Reform. *Phi Delta Kappan*, 83(1), 62-66.
- Goertz, M.E. (2005). Implementing the No Child Left Behind Act: Challenges for the states. *Peabody Journal of Education* 80(2), 73–89.
- Goertz, M.E., Floden, R.E., & O'Day, J. (1995). *Studies of education reform: Systemic reform*. New Brunswick: Consortium for Policy Research in Education.
- Goldhaber, D., & Hannaway, J. (2004). Accountability with a kicker: Observations on the Florida A+ accountability plan. *Phi Delta Kappan*, 85(8), 598-605.
- Goldstein, L. (2008). Kindergarten teachers making "street-level" education policy in the wake of No Child Left Behind. *Early Education and Development*, 19(3), 448-478.
- Goodwin, R.H., Cunningham, M.L., & Eagle, T. (2005). The changing role of the secondary principal in the United States: An historical perspective. *Journal of Educational Administration and History*, 37(1), 1-17.
- Grant, G. (1983). The teacher's predicament. *Teachers College Record*, 84(3), 593-609.
- Guarino, C., Santibanez, L., & Daley, G. (2006). Teacher recruitment and retention: A review of the recent empirical literature. *Review of Educational Research*, 76(2), 173-208.
- Guggino, P., & Brint, S. (2010). Does the No Child Left Behind Act help or hinder K-12 education? *Policy Matters*, 3(3), 1-7.
- Hamilton, L.S. (2003). Assessment of a policy tool. *Review of Research in Education*, 27, 25-68.
- Hamilton, L.S., Stecher, B., Marsch, J., McCombs, J., Robyn, A., Russell, J. (2007). *Standards based accountability under no child left behind: Experiences of teachers in three states*. Santa Monica, CA: Rand Corporation.
- Hamilton, L.S., Stecher, B., Yuan, B. (2008). *Standards-based reform: History, research, and future directions*. Santa Monica: Rand Corporation.

- Hammond, B. (2009). No Child Left Behind law sanctions Oregon school at record level, *The Oregonian*.
- Hannaway, J., & Woodroffe, N. (2003). Policy instruments in education. *Review of Research in Education*, 27(2003), 1-24.
- Hanson, E.M. (2006). *Educational administration and organizational behavior* (5th ed.). Boston: Allyn and Bacon.
- Hanushek, E.A., Kain, J.F., & Rivkin, S.G. (2004). Why public schools lose teachers. *Journal of Human Resources*, 39(2), 326-354.
- Hanushek, E.A., & Raymond, M. (2002). *Lessons about the design of state accountability systems*. Paper presented at the Taking Account of Accountability: Assessing Policy and Politics, Harvard University.
- Heise, M. (1994). Goals 2000: Educate America Act: The federalization and legalization of educational policy. *Fordham Law Review*, 63, 345-381.
- Hill, P. (1999). Getting it right the eighth time: Reinventing the federal role. In M. Kanstoroom & C. E. J. Finn (Eds.), *New directions: Federal education policy in the twenty-first century* (pp. 147-170). Washington DC: Thomas B. Fordham Foundation.
- Hill, P., Cross, C., & Kilgore, S. (2000). The federal role in education. *Brookings Papers on Education Policy*(3), 11-57.
- Hirsch, E. (2005). Listening to the experts: A report on the 2004 South Carolina teacher working conditions survey.
http://www.teachingquality.org/pdfs/TWC_SCFinalReport.pdf
- Hollingsworth, S., Gallego, M.A., Clandinin, D.J., Morrell, P., Portes, P., Rueda, R. (2007). Editorial teams' introduction: Special issue on No Child Left Behind. *American Educational Research Journal*, 44(3), 454-459.
- Ingersoll, R. (1996). Teachers' decision-making power and school conflict. *Sociology of Education*, 69(April), 159-176.
- Ingersoll, R. (2003). *Who controls teachers' work? Power and accountability in America's schools*. Cambridge: Harvard University Press.
- Ingersoll, R. (2004). Why do high-poverty schools have difficulty staffing their classrooms with qualified teachers? *Renewing our schools, Securing our future* Pittsburgh: Center for American Progress.

- Ingersoll, R. (2007). Short on Power, Long on Responsibility. *Educational Leadership*, 65(1), 20-25.
- Ingersoll, R. (2012). Power, accountability, and the teacher quality problem. In S. Kelly (Ed.), *Assessing teacher quality* (pp. 97-110). NY, NY: Teachers College Press.
- Ingersoll, R., & Alsalam, N. (1997). Teacher Professionalization and Teacher Commitment: A Multilevel Analysis (NCES 97-069). Washington D.C.: National Center for Education Statistics.
- Ingersoll, R., & May, H. (2010). The magnitude, destinations, and determinants of mathematics and science teacher turnover. In CPRE (Ed.): Consortium for Policy Research in Education Research Report # RR-66.
- Ingersoll, R., & May, H. (2011). Recruitment, retention and the minority teacher shortage: Consortium for Policy Research in Education Report No. 69.
- Ingersoll, R., & Merrill, E. (2011). The status of teaching as a profession. In J. Ballantine & J. Spade (Eds.), *Schools and society: A sociological approach to education* (Vol. 4, pp. 185-198). CA: Pine Forge Press/SAGE Publications.
- Jennings, J., & Rentner, D. (2006). Ten big effects of the No Child Left Behind Act on public schools. *Phi Delta Kappan*, 88(2), 110-113.
- Johnson, S.M. (2006). The workplace matters: Teacher quality, retention, and effectiveness. Working paper. Washington D.C.: National Education Association.
- Kafka, J. (2009). The Principalship in Historical Perspective. *Peabody Journal of Education*, 84(3), 318-330.
- Kanstoroom, M., & Finn, C.E.J. (Eds.). (1999). *Overview: Thirty-four years of dashed hopes*. Washington DC: Thomas Fordham Foundation.
- Kantor, H., & Lowe, R. (2006). From New Deal to no deal: No Child Left Behind and the devolution of responsibility for equal opportunity. *Harvard Educational Review*, 76(4), 474-502.
- Karen, D. (2005). No Child Left Behind? Sociology ignored! *Sociology of Education*, 78(April), 165-182.
- Kim, I., & Loadman, W. (1994). Predicting teacher job satisfaction. Columbus, Ohio State University: ERIC Document Reproduction Service No. ED282707.
- Kingdon, J. (1995). *Agendas, alternatives, and public policies*. New York: Longman.

- Leiter, J. (1981). Perceived teacher autonomy and the meaning of organizational control. *The Sociological Quarterly* 22(2), 225-239.
- Lieberman, A. (1982). Practice makes policy: The tensions of school improvement. In A. Lieberman & M. McLaughlin (Eds.), *Policy making in education: Eight-first yearbook of the National Society for the Study of Education* (pp. 249-269). Chicago: University of Chicago Press.
- Linn, R. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4-16.
- Lipsky, M. (1980). *Street-level bureaucracy: Dilemmas of the individual in public services*. New York, NY: Russell Sage Foundation
- Loucks, H. (2007). A teachers' union perspective on NCLB accountability issues. *Pearson School, After Student Standards: Alignment*(2007), 199-211.
http://www.pearsonschool.com/live/assets/200727/2005_15Loucks_550_1.pdf
- Luke, D. (2004). *Multilevel modeling*. Thousand Oaks, CA: Sage Publications.
- Lyons, J.E., & Algozzine, B. (2006). Perceptions of the impact of accountability on the role of principals. *Education Policy Analysis Archives* 14(16), 1-19.
- Malen, B., & Muncey, D. (2000). Creating "A new set of givens"? The impact of state activism on school autonomy. In N. Theobald & B. Malen (Eds.), *Balancing local control and state responsibility for K-12 education* (pp. 199-244). Larchmont, NY: Eye on Education.
- Malen, B., & Ogawa, R.T. (1988, Winter). Professional-patron influence on site-based governance councils: A confounding case study. *Educational Evaluation and Policy Analysis*, 10(4), 251-270.
- Manna, P. (2002). *Federalism, agenda setting, and the dynamics of federal education policy*. Paper presented at the Annual Meeting of the Political Science Association, Boston, MA.
- Manna, P. (2006a). Control, persuasion, and educational accountability : Implementing the No Child Left Behind Act. *Educational Policy*, 20(3), 471-494.
- Manna, P. (2006b). *School's in: Federalism and the national education agenda*. Washington, D.C.: Georgetown University Press
- Margolis, J. (2008). What will keep today's teachers teaching? Looking for a hook as a new career cycle emerges. *Teachers College Record*, 110(1), 160-194.
- Marx, R., & Harris, C. (2006). No Child Left Behind and science education: Opportunities,

- challenges, and risks. *The Elementary School Journal*, 106(5), 467-478.
- Massell, D. (1998). State strategies for building local capacity: Addressing the needs of standards-based reform. CPRE Policy Briefs (pp. 3-15). Philadelphia: Center for Policy Research in Education.
- Mathison, S., & Freeman, M. (2003). Constraining elementary teachers' work: Dilemmas and paradoxes created by state mandated testing. *Education Policy Analysis Archives*, 11(34), 1-23.
- McDermott, K. (2003). What causes variation in states' accountability policies? *Peabody Journal of Education*, 78(4), 153.
- McDermott, K., & Jensen, L. (2005). Dubious sovereignty: Federal conditions of aid and the No Child Left Behind Act. *Peabody Journal of Education*, 80(2), 39-56.
- McDonnell, L.M. (2004). *Politics, persuasion, and educational testing*. Cambridge, MA: Harvard University Press.
- 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.
- McLaughlin, M., & Talbert, J. (2003). Reforming districts: How districts support reform. Seattle: Center for the Study of Teaching and Policy.
- McNeil, L.M. (1988). Contradictions of control, Part 2: Teachers, students and curriculum. *Phi Delta Kappan*, 69(6), 432-438.
- McNeil, M., & Klein, A. (2011, September 28). Obama outlines NCLB flexibility plan: Waives cornerstone provisions of Law, *Educatoin Week*.
- Meier, D., & Wood, G. (Eds.). (2004). *Many Children Left Behind: How the No Child Left Behind Act is Damaging Our Children and Our Schools*. Boston: Beacon Press.
- Minnici, A., Hill, D., & Kober, N. (2007). Educational architects: Do state education agencies have the tools necessary to implement NCLB? In C. o. E. Policy (Ed.), *From the capital to the classroom: Year 5 of the No Child Left Behind Act*. Washington D.C.: Center on Education Policy.
- Mintrop, H. (2003). The limits of sanctions in low-performing schools: A study of Maryland and Kentucky schools on probation *Educational Evaluation and Policy Analysis*, 11(3). <http://epaa.asu.edu/epaa/v11n3.html>.
- Mintrop, H., & Sunderman, G. (2009). *Why high stakes accountability sounds good but doesn't work: And why we keep on doing it anyway*. Los Angeles, CA: The Civil Rights Project/Proyecto Derecho Civiles at UCLA.

- O'Day, J. (2002). Complexity, accountability, and school improvement. *Harvard Educational Review*, 72(3), 1-31.
- Ogawa, R.T., Haymore Sandholz, J., Martinez-Flores, M., & Paredes Scribner, S. (2003). The substantive and symbolic consequences of a district's standards-based curriculum. *American Educational Research Journal*, 40(1), 147-176.
- Olsen, B. (2008). How reasons for entry into the profession illuminate teacher identity development. *Teacher Education Quarterly*, Summer (2008), 23-40.
- Opfer, V.D., Henry, G., & Mashburn, A. (2008). The district effect: Systemic responses to high stakes accountability policies in six southern states. *American Journal of Education*, 114(February 2008), 299-332.
- Paige, R. (2006). No Child Left Behind: The Ongoing Movement for Public Education Reform. *Harvard Educational Review*, 76(4), 461-473.
- Palmer, R.R. (1953). Is teaching a profession? *Phi Delta Kappan*, 34(4), 139-140, 142.
- Pearson, L.C. (1995). The prediction of teacher autonomy from a set of work-related and attitudinal variables. *Journal of Research and Development in Education*, 28(2 (Winter)), 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), 27-53.
- Pearson, L.C., & Moomaw, W. (2006). Continuing validation of the teaching autonomy scale. *Journal of Educational Research*, 100(11), 44-51.
- Pierson, P. (2005). The Study of Policy Development. *The Journal of Policy History*, 17(1), 34-51.
- Policy., C.o.E. (2012). Major accountability themes of second-round state applications for NCLB waivers. Washington, DC: Author.
- Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks: CA: Sage Publications.
- Retsinas, J. (1982). Teachers and professional authority. *Educational Forum*, 47(1), 25-36.

- Rice, J.K., Roellke, C., Sparks, D., & Kolbe, T. (2006). Hitting the target? Multi level case study findings from three states. Washington D.C.: Economic Policy Institute.
- Richer, S., Blanchard, C., & Vallerand, R. (2002). A motivational model of work turnover. *Journal of Applied Social Psychology*, 32(10), 2089-2113.
- Rorrer, A. (2004). Intersections in accountability reform: Complexity, local actors, legitimacy, and agendas. In L. Skrla & J. J. Scheurich (Eds.), *Educational equity and accountability: Paradigms, policies, and politics* (pp. 251-266). New York: Taylor & Francis Books.
- Rowan, B. (1990). Commitment and Control: Alternative strategies for the organizational design of schools. *Review of Research in Education*, 16(1), 353-389.
- Rudalevige, A. (2003). The politics of No Child Left Behind. *Education Next, Fall*, 63-69.
- Russell, J., & McCombs, J. (2006). Superintendent, principal, and teacher understanding of and attitudes about accountability. In R. Corporation (Ed.), (pp. 1-36). Santa Monica: RAND.
- Santoro, D.A. (2011). Good teaching in difficult times: Demoralization in the pursuit of good work. *American Journal of Education*, 118(1), 1-23.
- Settlage, J., & Meadows, L. (2002). Standards-based reform and its unintended consequences: Implications for science education within America's urban schools. *Journal of Research in Science Teaching*, 39(2), 114-127.
- Shipp, D. (2011). The politics of educational reform: Idea champions and policy windows. In D. Mitchell, R. L. Crowson & D. Shipp (Eds.), *Shaping education policy: Power and progress* (pp. 259-285). New York, NY: Routledge.
- Smith, M., O'Day, J., & Cohen, D. (1991). A national curriculum in the United States? *Educational Leadership*, 49(1), 74-81.
- Spillane, J., & Zeuli, J. (1999). Reform and teaching: Exploring patterns of practice in the context of national and state mathematics reforms. *Educational Evaluation and Policy Analysis*, 21(1), 1-27.
- Stecher, B., Epstein, S., Hamilton, L.S., Marsh, J., Robyn, A., McCombs, J. (2008). Pain and gain: Implementing No Child Left Behind in three states, 2004-2006. In RAND (Ed.). Santa Monica, CA.
- Stecher, B., & Hamilton, L.S. (2002). Putting theory to the test. *Rand Review, Spring 2002*, 17-23.

- Strike, K. (1990). Is teacher a profession: How would we know? *Journal of Personnel Evaluation in Education*, 4, 91-117.
- Sunderman, G. (2006). The unraveling of No Child Left Behind: How negotiated changes transform the law. Cambridge, MA: The Civil Rights Project at Harvard.
- Sunderman, G., & Kim, J. (2007). The Expansion of Federal Power and the Politics of Implementing the No Child Left Behind Act. *Teachers College Record*, 109(5), 1057-1085.
- Sunderman, G., & Orfield, G. (2006). Domesticating a revolution: No Child Left Behind reforms and state administrative response. *Harvard Educational Review*, 76(4), 526-556.
- Sunderman, G., Tracey, C., Kim, J., & Orfield, G. (2004). Listening to teachers: Classroom realities and No Child Left Behind. Cambridge, MA: The Civil Rights Project at Harvard University.
- Superfine, B.M. (2005). The politics of accountability: The rise and fall of Goals 2000. *American Journal of Education*, 112, 10-43.
- Sykes, G. (1999). The "New Professionalism" in education: An appraisal. In J. Murphy & K. Seashore-Louis (Eds.), *Handbook of research on educational administration (2nd edition)* (pp. 227-249). San Francisco: Jossey-Boss.
- Talbert, J., & McLaughlin, M. (1994). Teacher professionalism in local school contexts. *American Journal of Education*, 102(2), 123-153.
- Tourkin, S., Thomas, T., Swaim, N., Cox, S., Parmer, R., Jackson, B. (2010). Documentation for the 2007–08 Schools and Staffing Survey (NCES 2010-332). Washington, D.C.: U.S. Department of Education.
- Umpstead, R.R. (2008). The No Child Left Behind Act: Is it an unfunded mandate or a promotion of federal educational ideals? *Journal of Law and Education*, 37(2), 193-229.
- Usher, A. (2011). *AYP results for 2010-11*. Washington, D.C.: Center for Education Policy
- Valli, L., Croninger, R., Chambliss, M., Graeber, A., & Buese, D. (2008). *Test driven: High-stakes accountability in elementary schools*. New York: Teachers College.
- Valli, L., & Buese, D. (2007). The changing roles of teachers in an era of high-stakes accountability. *American Educational Research Journal*, 44(3), 519-558.

- Wanker, W.P., & Christie, K. (2005). State implementation of the No Child Left Behind Act. *Peabody Journal of Education*, 80(2), 57-72.
- Weatherly, R., & Lipsky, M. (1977). Street-level bureaucrats and institutional innovation: Implementing special-education reform. *Harvard Educational Review*, 47(2), 171-197.
- Week, E. (1999). Quality Counts 1999, *Education Week*.
- Weiss, E.M. (1999). Perceived workplace conditions and first-year teachers' morale, career choice commitment, and planned retention: A secondary analysis. *Teaching and Teacher Education*, 15, 861-879.
- White, P.A. (1992). Teacher empowerment under "ideal" school-site autonomy. *Educational Evaluation and Policy Analysis*, 14(1), 69-82.
- Williams, J. (2008). School districts face sanctions under landmark law, *USA Today*.
- Wong, K., & Sunderman, G. (2007). Education accountability as a presidential priority: No Child Left Behind and the Bush presidency. *Publius: The Journal of Federalism*, 37(3), 333-350.
- Wood, G. (2004). A view from the field: NCLB's effects on classrooms and schools. In D. Meier & G. Wood (Eds.), *Many children left behind: How the No Child Left Behind Act is damaging our children and our schools* (pp. 33-50). Boston: Beacon Press.