8-2016

Correlates and Effects of Data-Informed Decision-Making: An Empirical Examination of the Loose Coupling Theory Using a National Data Set

Xingyuan Gao
Western Michigan University, xingyuangao@gmail.com

CORRELATES AND EFFECTS OF DATA-INFORMED DECISION-MAKING: AN EMPIRICAL EXAMINATION OF THE LOOSE COUPLING THEORY USING A NATIONAL DATA SET

Xingyuan Gao, Ph.D.

Western Michigan University, 2016

The notion of schools as "loosely-coupled" organizations has been widely discussed in the research literature. Many argue that the loose coupling is either a protective mechanism for schools to buffer external pressure or a barrier to implementing new reforms. Against the backdrop of standards-based reform and accountability policies in the last 15 years, which focus on tightening the system as the way to improve K-12 schools, I conducted three related studies to empirically test the loose coupling theory by applying two-level hierarchical linear models to nationally representative data in the US.

In the first study I evaluated the loose coupling theory by examining the association between data-informed improvement efforts at the school level and data-informed instruction at the classroom level. Statistically significant associations were identified. However, the effect sizes were small, and a very small proportion of variance was explained. In the second study I evaluated the loose coupling theory through examining the association between data-informed improvement efforts at the district and school levels. Statistically significant associations were identified with large effect sizes, and a large proportion of variance was explained. In the third study, I examined whether school performance was attributable to data-informed decision-making at the district, school, and classroom levels. Only one school level data-informed decision-making area was statistically significant.

In summary, viewing "using data" as a coupling mechanism and "decision-making" as a coupling element, I found that the K-12 system appears to be "loosely-coupled" between the school and classroom levels, but "tightly-coupled" between the district and school levels. In other words, the blanket statement of "schools as loosely coupled organizations" is not accurate. Thus, the dominant reform agenda in the last 15 years based on this blanket statement—advocating to tighten the system via curriculum standards, accountability tests, and evaluation as the way to improve the K-12 schools—is called into question. Policy implications are discussed.

CORRELATES AND EFFECTS OF DATA-INFORMED DECISION-MAKING: AN EMPIRICAL EXAMINATION OF THE LOOSE COUPLING THEORY USING A NATIONAL DATA SET

by

Xingyuan Gao

A dissertation submitted to the Graduate College
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
Educational Leadership, Research and Technology
Western Michigan University
August 2016

Doctoral Committee:

Jianping Shen, Ph.D., Chair
Sue Poppink, Ph.D.
Yuanlong Liu, Ph.D.

ACKNOWLEDGMENTS

I would never have been able to finish my dissertation without the guidance of my committee members, help from my professors, and support from my friends and family.

I would like to extend deepest thanks and appreciation to my dissertation mentor, committee chair and academic advisor, Dr. Jianping Shen, for his excellent guidance, thought-provoking suggestions and encouragement throughout the process. I am very grateful for the expert training he offered me all the way from the very beginning of scholarship in educational leadership through to the completion of my dissertation. Dr. Shen provided extreme patience in the face of numerous obstacles. I am truly fortunate to have had the opportunity to work with him. Without Dr. Shen, the successful completion of my dissertation would not have been possible. Sincere thanks also go to my committee members, Dr. Sue Poppink and Dr. Yuanlong Liu, for their participation and insightful words of advice on research and U.S education policy and law, along the way.

I would also like to thank all the professors for their friendly guidance and input on my dissertation. I am especially thankful to Dr. Jessaca Spybrook for her professional courses and helpful recommendations on my research methodologies.

I would also like to thank all the staff at WMU, who provided support and assistance. I extend additional appreciation to Chair Dr. Donna Talbot for her support. In a similar vein, special thanks also go to Diane Bourgeois and Lynda Root, for their wonderful support.

I would like to thank all the staff at Department of Education at East China Normal University. They always supported and encouraged me to pursue my dream. Without the encouragement from professors Chengxian Du, Bin Zhou, and others, I would not have had the opportunity to study at WMU and finish my dissertation.

I would like to thank Dr. Jiangang Xia for his supports and advices on methodology. I would like to thank Harold Curtis, for his help on my English writing and friendship.

Last, but certainly not least, I would especially like to whole heartedly thank my family. I am so grateful for my parents, Shengchang Gao and Guifang Chen, for their unwavering support and love with their best wishes. I would also like to extend my thankfulness to my parents-in-law, Guoqing Gu and Boyan Huang, for their help and share of household burdens for my family. I would like to give special thanks to my wife, Yuwen Gu, who sacrificed a vast amount of time in shouldering far more than her fair share of parenting. And many thanks to my precious daughters, Mina Gao and Hannah Gao. They always believe in me, support me and love me, despite my many hours of distraction from family life over the past years.

TABLE OF CONTENTS

Table of Contents—Continued

CHAPTER

Table of Contents—Continued

CHAPTER

Table of Contents—Continued

CHAPTER

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Educational reform has been one of the most important topics in the United States in the past century. The public's growing dissatisfaction with the condition of education has urged waves of reform endeavors to overhaul educational governance structure (Fusarelli, 2002). However, it appears that educational organizations are incorrigibly obstinate—educational organizations, from districts, schools to classrooms, were hardly changed from past waves of reform, especially at the school's technical core (see Goodlad & Klein, 1975; D. Cohen, 1990; Sarason, 1990, etc.). Why is it so difficult to make changes in educational organizations? Organizational theorists have provided different answers, including organizational inertia (see Hannan & Freeman, 1984), isomorphism (see DiMaggio & Powell, 1991), complex theory (see Griffiths, Hart, & Blair, 1991; O'Day, 2002), and so on. Among those important theories, there is one from Weick's (1976) seminal paper: *educational organizations as loosely coupled systems*, which argued that educational organizations were essentially loosely coupled.

The U.S. Education has witnessed increasing emphasis on tightening schools by standardization and accountability. During the past two decades, federal and state government gradually evolved as the most important players in reforming of public education. The No Child Left Behind Act (NCLB) in 2002 highlighted the responsibility of states, districts, and schools to use data for school evaluation and improvement (Mandinach, Honey & Light, 2006). In 2009, the Race to the Top (RttT), as part of the American Recovery and Reinvestment Act, encouraged states to establish aligned assessments and to create longitudinal data systems for evidence-based instructional improvement (Russell, Meredith, Childs, Stein, & Prine, 2015). Using data to

inform decisions has emerged as a critical strategy for not only accountability purposes in public education (Bernhardt, 2013; Coburn & Turner, 2012; Mandinach, 2012), but also for continuous improvement for transforming schools into professional learning organizations (Fullan, 2007; McLaughlin & Talbert, 2006).

In this study, Weick's argument that "educational organizations as loosely coupled systems" is quantitatively evaluated from a new perspective—data using patterns across educational systems.

Problem Statement

Why have previous reform efforts had so little effect in changing the technical core of education—curriculum and instruction? The loose coupling theory suggests that the loose coupling status of schools is a major reason. In order to eliminate the loose coupling, federal and state government have promoted standards-based reform and performance-based accountability movement during the past two decades—both of which aimed to fundamentally change U.S. education by tightening the entire system. Nevertheless, researchers and practitioners criticized that neither reforms achieved their intended goals.

There could be a different perspective to guide reform. Most loose coupling theorists regard the loose coupling as the nature of educational organizations. However, the constructivism view of the theory makes it compatible with contemporary theories of organizational learning.

Data has played a critical role after NCLB, yet mostly for the accountability purpose. The potential of using data for continuous school improvement and organizational learning was underestimated. Also, while there are a great deal of studies reported the incoherent data using pattern among different levels of educational system, few of them take the loose coupling theory perspective.

Broader Problem

The broader problem this study is trying to address is the re-working of the loose coupling theory in the age of data.

A theory in crisis. It has been more than three decades since Weick (1976) published his masterpiece of organizational analysis— "educational organizations as loosely coupled systems". The loose coupling theory thrived during the 80s and 90s. It served as a very powerful theoretical tool for researchers and policy makers to understand why past educational reform often failed—most reform efforts were buffered by the loose coupling nature of educational organizations. Loose coupling keeps schools stable, but also makes them stubborn. Based on the theory, scholars began to discover and expound various phenomena in school settings that impeded educational reform—for example, it was found that little consensus existed among teachers from same schools, as well as principals from same districts (Deal & Celotti, 1980); inspections and tests were rarely used for the purpose of evaluation for teachers or principals (Cuban, 1990); and organizational goals were unclear and collegial supports were limited (Darling-Hammond & McLaughlin, 1995).

However, it appeared that loose coupling theory was gradually left behind by other theories such as new managerialism, or complex theory. Rowan (2002), one of the important writers on loose coupling for more than 25 years, admitted that "education systems have become far more rationalized and bureaucratized than they were when ideas about loose coupling first emerged in organization theory" (p. 604). He argued that the theory needs some substantial re-works to further serve organizational analysis in the future.

According to Rowan (2002), one of the major problems about loose coupling research is that studies are often careless in clarifying the "coupled elements" and "coupling mechanisms."

Likewise, H. D. Meyer (2002a) criticized some loose coupling studies, because they rarely referred to "any kind of linkage between any kind of organizational 'element', from schedules, routines, or policies, to attitudes, norms, and preferences" (p. 536).

Tightening of schools, a dead-end? Realizing that efforts to make educational changes were buffered by loose coupling nature of school, the policy makers straightforwardly took the "tightening" scheme. The U.S. education unprecedentedly witnessed a few major reforms which aimed to tighten the entire education system during the past decades. The state-led standards-based reform was introduced to form alignment among all educational elements, from curriculum, instruction, and text materials to testing. Nevertheless, research on standards-based reform suggested that effects of the alignment efforts were diminished by each level of the system (D. Cohen, 1995), thereby, the reform efforts could hardly reach the technical core of schooling (Shen & Ma, 2006). As Elmore (2000) argued: standards-based reform violates the fundamental premise of loose coupling—buffering the technical core from interference by external forces (p.8).

The next tightening scheme was the implementation of performance-based accountability systems in the entire nation. The accountability movement aimed to strengthen the effectiveness of standards-based reform by accountability measures based on mandatory high-stake standardized tests (Mason, 2003). The top-down accountability reform received even more criticism from both research communities and practitioners. It was found that accountability systems brought lots of negative impact on both teachers and students (see Desimone, 2013; Fitchett et al., 2014; Krieg, 2008; Sheldon & Biddle, 1998).

The criticism on standards-based reform and accountability reform called for more attention to studying the loose coupling nature of educational organization. First, both reforms

were designed to reduce variability and to remove misalignment within education systems, they deemed loose-coupling as an obstacle to change. Loose coupling theory regards loose coupling as a nature of education system that could be taken advantage of. Second, these "mainstream" reforms relied heavily on the top-down, external policies. Therefore, internal motivation to reform was suppressed. By contrast, the loose coupling theory suggests that loose-coupling can protect internal motivation by buffering external force. Lastly, it is very questionable that external forces can really "tighten up" the system, if so, why have teachers complained that their principals are becoming more invisible and disconnected to them after NCLB (Hargreaves & Goodson, 2006). Obviously, the previous reform attempts are not balanced and need significant improvement.

Researchable Problem

Ambiguity used to be the word to describe decision making in educational organizations a few decades ago. In the 1970s, researchers found decision making in a university was irrational and random (M. Cohen, et al., 1972), the process for decision making was simply "garbage in and garbage out". Similar patterns were pervasive in most education enterprises. It was reported schools were rarely evaluated with student achievement data by districts (J. W. Meyer & Rowan, 1978), and productivity of teachers was hardly ever measured (Cuban, 1990). Loose coupling theorists noted that the ambiguity of decision making reinforced the autonomy of information gatherers, and consequently a loosely coupled system (March, 1987; Orton & Weick, 1990).

Policy makers have determined to remove the ambiguity in education system. Since the enforcement of NCLB, reliance on data to guide decision making has become an important principle to eliminate the ambiguity of decision making in educational organizations. Accountability system in public education generate a large amount of data every day.

Consequently, a circle of continuous improvement was proposed within the education system—teachers use data to improve their instruction and assess students; Schools and districts use data to inform important decisions, including school plans, curriculum evaluation, teacher evaluations, and so on.

Such a circle of improvement essentially introduced a new plausible coupling mechanism to education system—educational decision makings within the organizations are now all connected with data. In another words, the coupling element is "decision-making", and the coupling mechanism is "using data". A researchable problem is revealed: based on the loose coupling theory, has the introduction of data application became a new coupling mechanism to link practitioners' decision-making across all levels of education system?

Studies Addressing the Problem

Studies on Coupling Mechanism within Educational Organizations

Earlier loose coupling studies were interested in the authority of office and task interdependencies, which were the perspectives featured in Weick's original discussion (Rowan, 2002). Ingersoll (1993; 1994; 1996) conducted a series of studies on power relationship in school. He argued that the educational organizations were loosely coupled depended on the domains of control. Deal and Celotti (1980) reported that the district, school, and classroom operated independently, as were individuals within each level. Shen and Ma (2006) investigated the relationship among the state's guidelines, school's curriculum, classroom teacher's instruction. They found that the influence of states and schools stopped at the classroom door.

More recent loose coupling studies incorporated some new managerial concepts, such as "symbols, interests, contracts, information, resource flows" (Rowan, 2002, p. 60). Spillane,

Parise, and Sherer (2011) discussed re-designing organizational routines to change instructional practice of teachers. Kreysing (2002) illustrated utilization of contract management to promote "controlled autonomy" in universities. Young (2006) suggested that alignment in agenda-setting functions and collaborative norms could enhance the coupling between teacher's instruction and leadership's expectation. P. E. Johnson and Chrispeels (2010) identified four types of linkages in the education systems, including (a) ideological linkage, (b) relational and communication linkages, (c) structural linkages, and (d) resource linkages. Spillane and Burch (2003) highlighted the effect of sense-making in influencing teachers' implementation of policy. Hökkä and Vähäsantanen (2013) proposed that coupling mechanisms should be teacher-centre to achieve a balance between loose coupling and tight coupling.

Studies on Data Using Patterns

There are many studies which have documented how districts could shape data-informed decision-making at school and classroom level (e.g., Anderson, Leithwood, & Strauss, 2010; Togneri & Anderson, 2003). However, more studies found that teachers, schools and districts tended to have differences in data collecting (Fuchs, 2004; Honig & Coburn, 2008; Shen, et al, 2010); data interpretation (Ingram, Louis, & Schroeder, 2004; Lachat, Williams & Smith, 2006; Rudner & Boston, 2003); and data utilization (Diamond & Cooper, 2007; Hess & Fullerton, 2009; Ikemoto & Marsh, 2007; Little, 2012). Those differences of data using pattern across the education system could be attributed to different mental models (Coburn & Talbert, 2006; Datnow & Hubbard, 2015; Jimerson, 2014), practitioner's incapability of using data (Dennis, 2008; Mason, 2002; McNamara, 2000; Reeves & Burt, 2006), insufficient support of data coaching (Honig & Hatch, 2004; Honig, 2006; Honig, 2012), and lack of a culture of using data (Datnow, Park, & Wohlstetter, 2007; Thornton & Perreault, 2002)

Significance of the Study

I agree with Rowan's (2002) acknowledgement that the loose coupling theory needs "substantial re-work", yet the theory has great potential to guide more successful school reform by connecting to organizational learning theories. The present study investigates how "using data" as a coupling mechanism and "decision making" as a coupling element are associated across classroom, school, and district levels. The study tries to fill a void in both loose coupling literature and educational data using literature, and may benefit researchers, policy makers, and practitioners.

For researchers and organizational theorists, this study may benefit the re-work of the loose coupling theory in three aspects. First, the present study would update the classic loose coupling theory within a more modern school context, where data is the core of schooling, and continuous improvement is becoming a new form of educational reform. Second, the study would draw researchers' attention from external incentives to internal motivation, and from top-down managerialism to bottom-up professionalism. Although it is true that external incentive and top-down managerialism play an indispensable role in organizational change, it is also true that long-lasting, sustainable changes need to be developed from inside, especially in the field of education. Lastly, the current study would address a typical problem of loose coupling studies. As Rowan (2002) asked: "where are empirical examples designed to measure the 'strength' of couplings among system elements, especially in large-scale systems where such couplings cannot be observed easily except during periods of massive disruptions to equilibrium?" (p. 607). This study is responding to Rowan's question—the empirical study presented here evaluates the strength of "using data" (coupling mechanism) on "decision making" (elements), with a large-scale, post-NCLB national-level data set. This is rare in existing loose coupling study.

For policy makers, the loose coupling theory provides a complete different angle to design reform. The loose coupling theory views loose coupling as a nature of educational organization, rather than a problem to be eliminated. Goldspink (2007) asserted that the loose coupling theory is founded on "a constructivist epistemology, making it more compatible with contemporary theories of learning and therefore educational reform which draws on or embraces constructivism" (p. 40). To some extent, loose coupling theorists shared some common grounds with organizational learning theorists. Specifically, this current study will discuss the role of data-informed decision-making in the combination of the loose coupling theory and professional learning organization, and contribute new knowledge to both models.

For education practitioners, this study may also be practically useful. Education practices in the real world can be much more complicated than the models in the theories. Therefore, educational practices need more detailed guidelines than magnificent blueprints. After decades of development, the loose coupling theory has switched from ideology debates (e.g., loose vs. tight) to more precise understanding and sorting of specific mechanism of coupling. Meanwhile, current knowledge about what can facilitate teacher adoption of data practice is very limited (Dunn, Airola, Lo, & Garrison, 2013). Few studies put the discussion of data practice within the context of organizational change. The relationship between teacher data using and school data using, and the relationship between school data using and district data using remained unclear. This study is designed to investigate the relationship based on the real data from a national survey of education practice of teachers, principals, and central office staffs, which may fill a void in literature and provide guidance for practitioners.

Purpose Statement

The U.S. government has shown continuing interest in pushing educational changes through tightening the loosely-coupled public school system during the past two decades. However, sufficient literature has suggested the current tightening strategy is not perfectly working—in terms of positively impacting the technical core of education.

I use post-NCLB national representative data to test whether the data-informed school improvement efforts are associated with teachers' patterns of incorporating data into instruction, as well as whether district-level data-informed improvement efforts can match with school-level data using practice. If association is found, I may argue that the new coupling mechanism—"using data" has been a way to penetrate the school gates and classroom doors to change the classroom-level instructional decision making. If there is no association, then "using data" has not yet been a new mechanism to impact practice in schools and classroom.

Conceptual Framework

The study is investigating how the practice of data using pattern at district, school, and classroom levels are associated, in the context of organizational reform which aims to transcend the looseness of the school system.

A graphic conceptual framework (Figure 1) has been developed to examine correlates of data patterns and other related organizational reform conceptions. The conceptual framework indicates that (a) teachers' data-informed decision-making at the classroom level is a function of schools' data-informed decision-making, (b) schools' data-informed decision-making is, in turn, a function of district's data-informed decision-making.

The notion of schools as "loosely coupled" organizations has been widely discussed in the research literature. Many argue it is either a protective mechanism for schools to buffer

external pressure or a barrier for implementing new reforms. Against the backdrop of systemic change and accountability, I applied two two-level hierarchical linear models to nationally representative data in the US, testing the "loosely coupled" theory through examining the association between data-informed improvement efforts at the school level and data-informed instruction at the classroom level, as well as the association between data-informed decision-making improvement efforts at the district level and data-informed improvement efforts at the school level. In addition, I also studied whether data-informed decision-making efforts at various levels are associated with school-level student achievement.

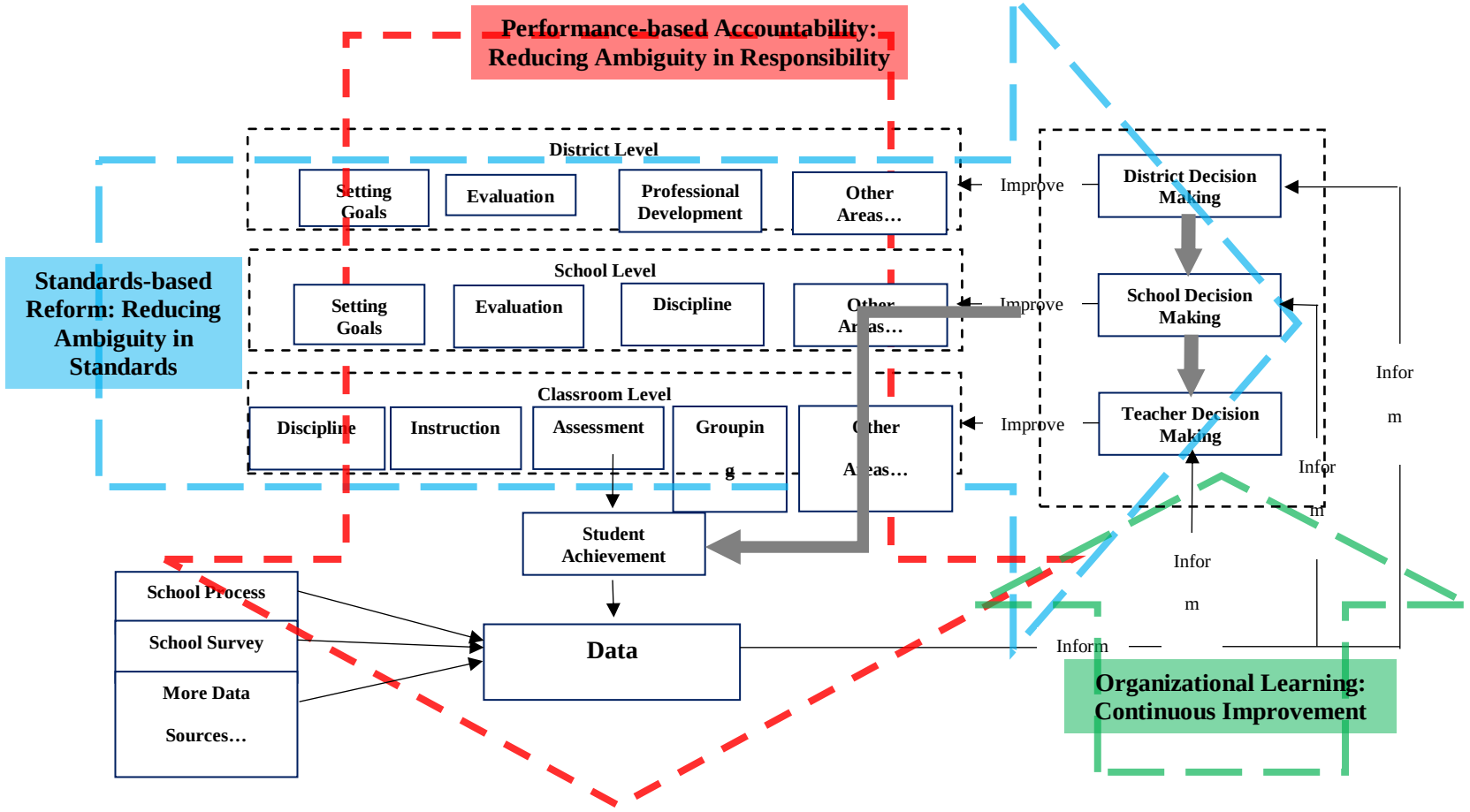


Figure 1. Conceptual Framework

As displayed in the conceptual framework, the three arrows in different colors represent three phrases of reform design—standards-based reform, performance-based accountability, and organizational learning. The blue "standards-based reform" arrow is horizontal, since it is not necessarily a top-down scheme. It aims to align educational elements (e.g., all educational activities, including "decision making" across all levels) to reduce variation within the system. However, the effect of blue arrow need to be reinforced by the red "performance-based accountability" arrow, which is a typical top-down strategy that runs through all three levels—from districts to schools, and to classrooms. The "data" box, which is at the tip of the red arrow, is largely composed of student achievement data. The "data" box goes to three "decision making" boxes via the links of "inform", and three "decision making" boxes connects to three levels of education system via the function of "improve". The big green "continuous improvement" arrow, implies a bottom-up reform strategy to initiate the process of data-informed decision-making. The logics, influence, and relationship of three reform arrows is to be discussed in the following chapters.

The foci of the study, however, are the three gray arrows from "district decision making" to "school decision making", from "district decision making" to "school decision making", and "district/school/teacher decision making" to "student achievement," respectively. They represent the current limited understanding of the strength of coherence among data-informed decision-makings at three organizational levels, and the effect of data-informed decision-makings from three organizational levels on student performance.

Research Questions

There are three studies for the dissertation. The first two studies are to investigate the relationship of data-informed decision-making pattern between schools and teachers, and

between districts and schools. The last study is designed to evaluate the influence of data-informed decision-making in terms of school performance.

The three research questions in the first study are:

(a) To what extent is teachers' data-informed decision-making attributable to the school level? (b) Is there a significant variation among schools in teachers' data-informed decision-making? And (c) after controlling for teachers', principals', and schools' background variables, are principals' and schools' data-informed decision-making associated with teachers' data-informed decision-making?

Similar three research questions guided the second study in terms of the data using pattern between districts and schools:

(a) To what extent is schools' data-informed decision-making attributable to the district level? (b) Is there a significant variation among districts in schools' data-informed decision-making? And (c) after controlling for principals', schools', and districts' background variables, are districts' data-informed decision-making associated with schools' data-informed decision-making?

The third study is guided by one research question:

To what extent is school performance attributable to data-informed decision-making at three levels?

Methods Overview

The data for this study is from a nationally representative School and Staffing Survey (SASS) 2003-04, which was designed by National Center for Education Statistics (NCES) of the U.S. Department of Education. Specifically, the SASS 2003-04 Public Teacher, Principal, School, and District were used. The study employed hierarchical modeling for the reason that

teachers are nested within schools and schools are nested within districts. Items from SASS database provide a rich set of variables, allowing control of teacher, school, and district characteristics when building the model.

Summary

The aim of the first chapter is to provide an overview of the present study. The chapter began with a brief review of loose coupling theory and previous reform efforts. It was argued that loose coupling theory could improve current reform by providing a different angle—a bottom-up approach with a focus on organizational learning, rather than the top-down approach with a focus on accountability. Data-informed decision-making, an emerging critical tool in education, could play a very important role in the shift of reform paradigm.

Data-informed decision-making is introduced to the education system at different levels, more educational stakeholders are required to use data to formulate decisions. Given the scenario that districts, schools, and teachers are required to make decisions based on data instead of subjective judgment, we proposed that it is possible the looseness of education systems be transcended through the process of using data across levels. I assume that "using data" has become a new coupling mechanism to connect the elements of "decision making" in school systems.

The statement of purpose, three groups of research questions, and the methods overview are presented in this chapter. For details of the literature review and methodology, please refer to Chapter Two and Chapter Three. Results and discussion are presented in Chapter Four and Chapter Five, respectively.

CHAPTER II

LITERATURE REVIEW

This review of literature contains four major sections, which are "the myth and ceremony of loose coupling in educational organizations", "transcending technical core—managerialism in education", and "the rise of data-informed decision-making", and "from loosely coupled organizations to organizational learning". After reviewing the literature of these three topics, a summary section was developed to show readers how these three streams of literature were intertwined, and how this literature might guide the following sections of the study.

Myth and Ceremony of Loose Coupling in Educational Organizations

"Education organizations as loosely coupled systems (Weick, 1976)" has been a long-discussed theory. The loose coupling theory has gained its popularity for its powerfulness in explaining various phenomenon in school systems in 70s. It attracted continuous attention as well as controversy from scholarship for almost half centuries.

Weick was the one of the researchers who provided the most comprehensive argument for the concept of "loose coupling" (Firestone & Wilson, 1985; Fusarelli, 2002). "Coupling" referred to linkages between different organizational elements (Orton & Weick, 1990), and "Loose coupling," as Weick (1976) originally proposed, meant that events were attached to each other at some degree, however, each still retained its identity. He noted that principals and teachers were connected, but the attachment was "circumscribed, infrequent, weak in its mutual affects, unimportant, and/or slow to respond" (p. 3).

Loose-coupling Phenomenon in Educational Organizations

Existing studies have documented loose-coupling phenomena in both school/classroom level, and district/school level.

Loose-coupling phenomena at school/classroom level. Weick's argument on "education organizations as loosely coupled organizations" was evidenced by a great deal of studies. In most cases, loose-coupling was found at school/classroom level, suggesting that individual teachers' work was essentially isolated from school administrations as well as other teachers' work.

Classrooms are the central arena in the school system (Deal & Celotti, 1980). Katz (1964) noted that autonomy in schools was inevitable as it could accomplish various functional requirements within the organization, but it was questionable that "how much autonomy must schools have to be effective as teaching centers" (p. 454). Goodlad and Klein's (1975) book, *Looking Behind the Classroom Door,* documented how teachers continued their practice regardless of the rhetoric of the reforms. D. Cohen (1990) and Spillane (1999) also revealed the difficulty of changing teachers' instructional practice. D. Cohen argued that teachers' resistance to external reform was not the only reason for failed educational reform. Instead, as D. Cohen noted, teachers did not have enough support for them to understand new instructional policies and to evaluate their classes against policies, which resulted in teachers' failure to implement reformed curriculum as states intended. Spillane compared teachers who changed their instruction to teachers who did not, and proposed the concept of "teachers' enactment zone". The enactment zone, according to Spillane, was a space that teachers "notice, construe, construct and operationalize the instructional ideas advocated by reformers" (1999, p. 144). Spillane (1999) noted three key issues related to the enactment zone that could result in the failure of changing teacher's core practice. These three issues were: a) the zones were individualistic instead of

social; b) the zones lacked rich deliberations with other teachers and reform experts; and c) the above mentioned deliberations had limited material resources or artifacts to support them. This literature by Goodlad, D. Cohen, and Spillane implied that it was very hard to make changes to the technical core of teachers if the improvement efforts only focused on specific areas or some components at a time instead of looking at the whole system with several levels.

Loose-coupling phenomenon at district/school level. While most researchers agreed that the loose-coupling phenomenon existed at the school/classroom level, it seemed there was less consensus on whether the loose-coupling existed at the district/school settings. Goldring and Hallinger (1992) defined three types of district control context: centralized district control context, decentralized district control context, and fragmented centralization of control.

It was believed that strong bureaucratic control (Cuban, 1990; Hightower, 2002; March, 1978; O'Day, 2002), district leadership (Murphy & Datnow, 2003), culture norms (Firestone & Wilson, 1985; Firestone, 2009), and professional support (J. Fitzgerald, 1993; Quellmalz, Shields, & Knapp, 1995) can help to form a tight association between districts and schools. However, studies found that the connection between schools and districts could be just as loose as the connection between schools and classrooms. For example, in a three-year study conducted by scholars at Stanford, researchers investigated into perceptions of superintendents, school administrators, and teachers on various topics—from organizational related to instructional related issues. The results suggested that disagreements were very common among professionals in the education systems. On the one hand, there was little consensus among teachers from the same schools, nor among principals from the same districts. On the other hand, principals tended to disagree with their teachers, and superintendents tended to disagree with their principals. The study concluded that "the three levels of educational organizations—district, school, and

classroom—operate independently, as do individuals within each level" (Deal & Celotti, 1980, p. 471). Similar pattern was also reported in Chrispeels, Burke, Johnson, and Daly's (2008) study, which suggested misaligned mental models among superintendents, principals, and teachers. The movement of site-based management also contributed to the looseness at the district/school level. Stevenson and Schiller (1999) conducted a national longitudinal study of high schools in the U.S. From 1980 to 1993, they found that in the states that encouraged site-based management, central offices had less influence over schools in various decision making domains.

Examining loose-coupling phenomena with different organizational models. For centuries, organization theorists have developed various management models to understand and evaluate the nature of educational organizations. The diversity of theories of management and organization implied that there was not a single model or theory that could explain all aspects of educational organization. Tony Bush (2010), the author of *Theories of Educational Leadership & Management* classified most prevailing educational organizational theories into six classic models: formal model, collegial model, political model, subjective model, ambiguity model, and cultural model.

Although a formal model was the fundamental framework for understanding educational organizations in general, organization theorists argued that different models need to be used simultaneously to improve the understanding of different dimensions of school settings. In many cases, organizational theorists tend to use different models or metaphors to understand distinct but complementary facets of the organizational reality.

In the literature, the phenomenon of loose-coupling was found to be deeply rooted in an educational organization's bureaucracy, collegiality, politics, and culture. Therefore, loose coupling had been investigated with a myriad of regarding classic frameworks. For example, J.

W. Meyer and Rowan (1978) contended that educational organizations lacked close internal coordination in terms of both bureaucratic and collegial aspects, especially for instructional activity. They argued that "loose coupling" was a result of weak centralized governmental and professional controls. Among the literature, bureaucratic models, collegial models, political models, ambiguity models, and cultural models were mostly discussed.

Bureaucratic Model. Cuban (1990) summarized recurring issues in American history of educational reforms and inquired into the intractability of school technical core. Cuban argued that tightly coupled organizing vanished when it came to the classroom instruction, because the two bureaucratic tools of school administration—inspections and tests—did not actually function as expected. Inspections were considered as more ritual than tools for control, and tests could hardly measure teacher productivity.

Collegial Model. Lortie (1975) explained why instruction seemed to be an isolated activity in the collegial aspects—experience of teachers tended to be private instead of shared. Even when teachers were students, they began to form their private experience of teaching by watching how their teachers taught. However, some other studies attempted to explore the reasons why instruction seemed to be isolated activity in bureaucratic aspects. In addition, Rosenholtz (1989) noted that the collegial support could hardly be effective without clear goals to inform the direction to collegiality: "There is an organizational basis for directing behavior, for motivating behavior, for justifying behavior, and for evaluating behavior" (p. 360).

Political Model. Cuban (1990) revealed a bargain between teachers and administrators and policy makers. Schools need teachers' support to achieve desired student outcomes. In exchange, they protect teachers from real external scrutinize.

Culture Model. J. W. Meyer and Rowan (1978) highlighted a "logic of confidence" within educational systems: people at different levels have confidence in each other's work. For example, the community and the board have faith in the superintendent, the superintendent believes in the principal, and the principal trusts teachers. Bryk et al (2009) suggested such trust was a moral resource for school improvement. In the meantime, Firestone (2009) described how a typical district with loosely coupled culture looked like:

> A loosely coupled district lacks a shared vision for teaching and learning … These districts don't attempt to send a common message to schools. Departments in the district, such as the curriculum office and the Title I office, often send divergent messages to schools. These districts often adopt and drop improvement initiatives quickly with no effort to coordinate new programs (p. 672)

Ambiguity Model. M. Cohen and his colleagues (1972) investigated into the ambiguity of decision making process in educational organizations. They found a great amount of decision makings were irrationally and randomly made, where elements for rational decision making were missing. Firestone and Wilson (1985) listed three characteristics of teaching to explain the ambiguity relationship between teachers and principal.

> First, neither criteria for success nor means of achieving it are clear. Even when the results come out right, teachers find it hard to know if they can take credit. This ambiguity leads to the second problem vulnerability. Teachers are sensitive to infringements of their authority and autonomy from both the public and students. Third, teaching is a lonely occupation with little chance to talk about one's work with others who can appreciate what one has done (p. 20).

Loose-coupling, Raison d'etre of School Systems?

In the literature, there has been a consensus that school is a loosely coupled system. However, there are widely differing ideas on how to deal with the "looseness" in the context of educational improvement. Scholars in the 1970s tended to take "education organizations as

loosely coupled systems" as the *raison d'être* of the school system. In Weick's original argument on loose coupling, Weick (1976) summarized seven advantages for being loosely coupled: First, loose coupling could enable organizations to endure little changes in the environment, which might inspire continuous archaic traditions and innovative improvisations. Second, as sand was better than rock in displaying wind currents because "sand has more elements, more independence among the elements, and the elements are subject to a greater amount of external constraint" (p. 6), independent sensing elements exist in loose coupling might drive the systems respond to their environments better, which implied a special "sensitive sensing mechanism". Third, a loosely coupled system was good at dealing with localization issues. Given the relatively loose inter-dependency, elements in the system had more flexibility to meet the local needs, thus no additional adjustments would be addressed to the systems. The last advantage of the loosely coupled system was its creativity. Since the diversity of elements in a loosely coupled system were preserved better than in a tightly coupled system, the loosely coupled systems were gifted with more individual ideas, solutions and information. Fifth, given the nature of independent status of elements in a loosely coupled system, the spread of deterioration was prevented. Sixth, loosely coupled systems, especially in terms of schools, encouraged more self-efficacy to teaching staff. Seventh, since loosely coupled systems required less resource for facilitating coordination, the cost for running the system would be less expensive than running tightly coupled systems. Among those advantages, Weick's third point on localization drew attentions of many scholars. A body of studies investigated into more detailed process of localization—how individual teacher or teacher communities understood, adapted, and re-shaped state policy or district policy into local environment (e.g., Coburn, 2001; Spillane, 1999); and how district

mediated between state policy and school practice (e.g., Honig, 2009; Spillane, 1994; Spillane, 1998).

Logic of confidence. Institutionalists provided a reasonable logic to explain why an educational organization should "decouple" its formal structure from its technical activities and outcomes—to maintain its legitimacy (J. W. Meyer & Rowan, 1978). Educational organizations might be deemed ineffective or inconsistent from the economizing perspective. However, instructional efficiency did not equal good education. Only through avoiding external inspection and preserving themselves in isolated environments, educators could focus on their relationship with students without external disturbance. J. W. Meyer and Rowan (1978) defended the decoupling of school formal structure for four reasons: First, the avoidance of close inspection increased the commitments of teachers. Second, close inspection could destroy many of the critical values of education. Third, decoupling protected local schools from technical uncertainty that was created by state and federal. Local schools, in most cases, need to redesign the state and federal curriculum so as to meet the specific needs of local students. Fourth, decoupling protected schools from inconsistent and conflicting institutionalized rules in plural environments by buffering units from each other. The isolation and autonomy of subunits avoided conflict and inconsistency between different categories of professionals.

While J. W. Meyer and Rowan's logic of decoupling in educational organizations seemed plausible, it is still hard to image that an organization could survive without proper inspection across its hierarchical structure. J. W. Meyer and Rowan (1978) highlighted a unique function of "confidence" within educational organizations. They stated that the secret rule under loosely coupled structure was the trust between educators or administrators at different levels. For example, the community and the board had faith in the superintendent, the superintendent

believed in his or her principals, and the principals trusted their teachers. However, none of these people really understand others' achievements. They just assumed others were doing a good job. Bryk et al. (2009) suggested that relational trust functions as "both a lubricant for organizational change and a moral resource for sustaining the hard work of local school improvement" (p. 207).

The advantage of teacher's confident feeling of being isolated and uninspected was confirmed by Deal and Celotti's (1980). Deal and Celotti reported:

> Attempting to tighten schools through competency tests, stringent teacher evaluation, higher levels of collaboration among teachers and specialists? may, by exposing inconsistencies and contradictions, endanger the existing relationship between communities and schools that is based mainly on confidence, trust, and belief (p. 472).

A double-edged sword. In Weick's (1976) classic argument on "educational organizations as loosely-coupled systems", he listed a few disadvantages of loosely coupled schools. As he noted, individual and organizational effectiveness could be moderated by structural looseness. When teachers were "enjoying" the freedom of controlling their work in the classroom, they were also isolated from supports from their colleagues and other resources from organizations. Consequently, "students can potentially exert more control over instructional activities than community expectations or the formal chain of command" (p. 7). Also, the loose coupling resulted in potential occurrence of conflicts. If there was little agreement on organizational rules and practices, skirmishes or clashes were almost inevitable among teachers and administrators. Furthermore, although a loosely coupled system had its advantage in that it could "isolate its trouble spots, and prevent the trouble from spreading" (p. 7), it would be a challenge to repair those trouble spots in the system. Lastly, even if the cost of running a loose-coupled system could be lower than the cost of running a tight-coupled one, fund allocation in loose-coupled systems was "unspecifiable, unmodifiable, and incapable of being used as means of change" (p. 8) since there were less rational decision makings in the systems.

The conflicts within educational organizations had been evidenced in many district settings as well. McLaughlin and Mitra (2001) reported that "the most dedicated teachers and principals will have a hard time sustaining reform practices and philosophy if their district context is hostile or pushing in incompatible directions" (p. 312). Leithwood et al. (2004) noted that "the chance of any reform improving student learning is remote unless district and school leaders agree with its purposes and appreciate what is required to make it work" (p. 7). Honig and Hatch (2004) revealed potential conflicts between schools' own strategy and external demand from district central offices. Spillane (2000) pointed out that central offices tend to focus on piecemeal changes and miss the whole picture of reform proposals.

Intractability of the "technical core". Technical core is generally considered as the central piece of "loose coupling" (e.g., Bryk, Sebring, Allensworth, Luppescu, & Easton, 2009; Elmore, 2000; J. W. Meyer & Rowan, 1978). Elmore (2000) reviewed previous work on loose coupling and offered a definition for "technical core":

> detailed decisions about what should be taught at any given time, how it should be taught, what students should be expected to learn at any given time, how they should be grouped within classrooms for purposes of instruction, what they should be required to do to demonstrate their knowledge, and, perhaps most importantly, how their learning should be evaluated—resides in individual classrooms, not in the organizations that surround them (p. 6).

Many scholars have written about the intractability of the "technical core." As O'Day (2002) postulated, teachers tended to "close" the classroom door simply because teachers' work was "subject to continual interruption as others try to thrust new information upon them" (p. 300), even worse, "much of the information was irrelevant to the improvement of instruction and learning" (p. 300). More recently, Bryk et al. (2009) noted that if the technical core is not well articulated for all teachers, they were inclined "to determine their own objective and enact instruction accordingly, leading to variation within the same school" (pp. 264-265).

Buffering effect. While J. W. Meyer and Rowan (1978) agreed that the "logic of confidence" in the loosely coupled educational systems brought some benefits, they admitted that the "logic of confidence" could also be problematic: the confidence placed at each level offered almost impermeable protections for their stable status and peaceful environment. Therefore, any endeavor to bring positive reform to the system was also buffered outside of schools or classrooms. On the one hand, loose coupling ensured the stability and durability of the systems. On the other hand, loose coupling made the system immune to any changes, even to good changes.

Unlike teachers, schools cannot simply close their doors to external controls. However, it was found that the seemly "overwhelming" bureaucratic control from central offices could not really improve the intractability of the technical core in loosely coupled schools. For example, Crowson and Morris's (1985) initial findings on principals' working hours seemed to support a tightly coupled district context— principals spent almost half of their working hours to responding to central offices. However, further analysis on principals' detailed activities revealed that merely 10% of principals' communication with districts was actually related to curriculum and instruction. Similarly, Hannaway and Sproull (1978) reported that districts' activities of management were not associated with school management processes in terms of improving student learning. As Elmore (1993) argued, key decisions of curriculum and teaching were "passed from states to districts, from districts to principals, and from principals to teachers, with little effective focus or guidance" (p. 116).

Volunteerism effect. Loose coupling not only buffered the technical core from the change efforts outside of the system, but also limited the collective or individuals' innovations within the system. Elmore (2000) discussed volunteerism effect in loosely coupled schools. He argued that

"volunteerism" had become the only way to improve practice in a loosely coupled system. Because principals and teachers were not inspected, they developed relative skills and knowledge to improve their leadership and instructional work only for the reasons of their own preferences and values, rather than for the reason of meeting some standards as conditions of their work. Elmore pointed out two issues of volunteerism:

> (1) innovations that are highly correlated with the personal values and predispositions of individual teachers and hence tend to be adopted only by a small proportion of receptive teachers at any given time; and,
>
> (2) innovations that are largely disconnected from any collective goal or purpose of the school or the school system. Schools are consequently almost always aboil with some kind of "change," but they are only rarely involved in any deliberate process of improvement, where progress is measured against a clearly specified instructional goal (p. 7).

Beyond Loose Coupling in Educational System

Existing knowledge informs us that loose coupling in educational organizations had both advantages and disadvantages—it provides the stability and durability to the systems, but it also diminishes reform and changes from both outside and inside. However, making changes to education systems is inevitable in the age of accountability. Under different paradigms? Both managerialism and professionalism claim to be able to improve the technical core, yet through different techniques. Supporters of managerialism advocate for tightening the system from the top-down, and establishing order and accountability. Supporters of professionalism, on the contrary, believe that "educational reform's progress depends on teachers' individual and collective capacity" (Stoll et al., 2006, p. 221). Therefore, teachers themselves were supposed to be the heroes who remove the intractability of the technical core. They argue that loose coupling is the nature of the educational organization -- to reach the technical core, educational improvement needs to be delivered from the bottom-up. While both camps have contributed to

the practice, two fundamental problems were frequently ignored in literature: which exact elements in the system should be coupled, and, through what mechanism?

Modifications to the concept of loose coupling. There had been some modifications to the blanket statement "educational organizations as loosely coupled systems". For example, Ingersoll (1993) argued that whether the educational organizations were loosely coupled depended on the domains of control. He suggested schools could be regarded as "tightly coupled" in certain domains, if researchers "look beyond oversimplified notions of bureaucracy and take into account a range of possible mechanisms of control" (p. 105). To illustrate this idea, Ingersoll (1994) examined the power domain of teachers and principals. He found that teachers had less control over social function of schooling, but had more control over classroom level activities. Another follow-up empirical study by Ingersoll (1996) showed that teachers lacked power over sorting and socialization functions in schools, but they were reported to have autonomy over instruction and curriculum in classroom. Ingersoll argued that a school was loosely coupled as far as the classroom was concerned, but was tightly controlled when it came to the school's sorting and socialization functions.

As the body of research on loose-coupling kept growing, Orton and Weick (1990) conducted an exhaustive review on 300 related studies to clarify Weick's points on loose-coupling. They summarized the utilizations of the concept of loose coupling into five categories: causation, typology, effects, compensations, and outcomes. Based on comparison of various applications of "loose coupling" in literature, Orton and Weick developed a comprehensive conceptualization for loose coupling. They cautioned that loose coupling was a dialectical concept rather than a unidimensional concept. They illustrated five kinds of possible misuse of concepts: first, applying research methodologies that encouraged researchers to parse dialectical

concepts into unidimensional variables; second, using it for flat, static descriptions, rather than detailed, dynamic descriptions; third, ignorance about the presence of connectedness within organizations; fourth, using loose-coupling as an imprecise general statements; and last, considering loose coupling as managerial failure (pp. 218-219). Although many studies conducted in 80s and 90s still regarded loose-coupling as an unidimensional concept, Orton and Weick's clarification was important to subsequent studies on loose-coupling, and was continually echoed by other researchers (e.g., P. E. Johnson & Chrispeels, 2010); Rowan, 2002; Meyer, 2002b; Spillane & Burch, 2003).

Coupling mechanisms in educational organizations. Clarifying the conception of loose coupling provided a basis for further work on identifying through which path certain elements in the system should be linked or remain loose. H. D. Meyer (2002a) pointed loose coupling literature rarely refers to "any kind of linkage between any kind of organizational 'element', from schedules, routines, or policies, to attitudes, norms, and preferences" (p. 536). Likewise, Rowan (2002) expressed his concern about existing research on loose coupling theory—studies were often careless in clarifying the coupling mechanisms in details.

Earlier studies on loose coupling used to focus on coupling mechanisms such as authority of office and task interdependencies, which featured in Weick's original discussion (Rowan, 2002). For example, Firestone and Wilson (1985) reported that through cultural linkages, principals could influence teachers' instruction. They posited that although bureaucratic linkages were necessary to provide restraints and opportunities for teachers, cultural linkages could serve to shape teachers' awareness about "how they think about what they do," including "individual's definition of the task" and "individual's commitment to the task" (p. 13). They further emphasized the interactions between bureaucratic linkages and cultural linkages, warning that if

the two linkages were not matched, there could be counterproductive results. Deal and Celotti (1980) observed that administrators could approach their subordinates through informal channels such as acting as "colleagues" or "symbolic leaders" to tighten the coupling of the educational system. Spillane and Burch (2003) argued that the coupling of administration and teaching varied by subjects and school level. Their interesting findings from 15 elementary school suggested that science was loosely coupled with administration and the policy environment, but language arts and mathematics were tightly coupled with administration and the policy environment.

More recent studies on loose coupling mechanism tended to incorporate newly developed concepts from the managerialism, Rowan (2002) noted that some newer studies turned to newer coupling mechanisms such as "symbols, interests, contracts, information, resource flows and so on" (p. 60). For example, Spillane, Parise, and Sherer (2011) suggested that by re-designing organizational routines, it is possible to regulate the instructional program, set and maintain direction, and transform classroom instruction. Kreysing (2002) explained how the implementation of contract management could promote the concept of "controlled autonomy", which allowed increased freedom of university faculties but maintained the influence of university central board. Spillane and Burch (2003) noted human agency and sense-making could influence teachers' implementation of policy. Young's (2006) case studies showed how district and school-level alignment in agenda-setting functions and collaborative norms influenced the coupling between teacher's instructional practice and leadership's expectation.

Striking a balance between loose coupling and tight coupling. Scholars begin to think whether it is possible to find a balance point between loose coupling and tight coupling. The key is to articulate and evaluate different types of coupling mechanism within the system. Meyer's (2002a) argument is enlightening to such attempts:

> …and to the counterintuitive fact that order does not necessarily result from suppressing or eliminating ambiguity and chance, but rather from artfully and nimbly harnessing or channeling them (p. 536).

P. E. Johnson and Chrispeels's (2010) study revealed various coupling linkages and clarified their functions within an education system. They analyzed qualitative data that was extracted from extensive interviews, focus groups, field note, document review, and observations of district leaders, principals, and school leadership teams. Their findings provided a map for future researchers to operationalize coupling mechanisms:

> The ideological linkage, addressing both end goals and means, provides the foundation on which to build school reform, enhance professional commitment, and improve student achievement. The ideological linkage defines the values, beliefs, mission, and core instructional principles, and gives form, coherence, and direction to the system. The relational and communication linkages pervade the system and determine whether it is possible to formulate a shared agreement (ideology) about the ends and means. The structural and resource linkages, if working well, provide the needed external directives and guidelines and corresponding internal policies and practices (pp. 770-771).

Teachers in Finland used to enjoy professional autonomy in classrooms. Now the tradition was also threatened by the increasing accountability pressure. Finnish scholars, (Hökkä & Vähäsantanen, 2013) proposed to achieve a balance between loose coupling and tight coupling via agency-centre coupling, where teacher professional agency is defined as teachers' ability to "negotiate the conditions and contents of their work, and to influence community and organizational issues, including educational reforms" (p. 134). Hökkä and Vähäsantanen (2013) argued that: because the fundamental coupling mechanisms were the actions and practices of individual teachers, the center of management practice should be teachers, their relationship, and their learning, instead of organizational structures and centrally determined standards. According to the authors, three elements were required to create an agency-centre coupling organization:

> (i) collaboration between actors within different levels of the organization,

(ii) high quality communication systems and communication and

(iii) shared meaning construction within and beyond organizational boundaries. (p. 148)

Transcending Technical Core—Managerialism in Education

Although some scholars in the 1970s and 80s tended to take "school as a loosely coupled organization" as a *raison d'être* of the school system (e.g., Deal & Celotti, 1980; J. W. Meyer & Rowan, 1978; Weick, 1976), since that time, others, such as those in favor of managerialism, have viewed tightening the loosely coupled education system as a mechanism for educational improvement. In the past decades, U.S. school systems have witnessed two important educational reform initiatives that were developed under the paradigm of new managerialism—performance-based accountability movement, and standards-based reform—both of which have changed the dynamics of the educational world (see Fusarelli, 2002; Meyer, 2002, etc).

New Managerialism—a Tightening Approach

Since precise technical requirements for instruction are difficult to define, local schools were able to buffer the influence from federal and states in the past decades (J. W. Meyer & Rowan, 1978; Pajak & Green, 2003). However, such practice of decentralization at local level is being suppressed by federal and states, according to a systemic review of education policy from 1950 to 2010 (Mitchell, 2011). The high-stake environment with constant emphasis on standardization and accountability has greatly changed the power relationship between local, state and federal levels of government (Malen, 2011). Through enactment of commands and mandates, the federal and state government gradually evolved as the leading force in education reform process. In contrast, the local community and parents are losing their power in school decision making. Given the growing demands for improved technical performance by governments, it becomes more and more impossible for local schools to ignore policy directives

by federal and states (Fusarelli, 2002). Loose coupling of school systems, which was once accepted by many researchers, is increasingly unbearable to the policy makers. Managerialism has become a mainstream thinking to guide the tightening of the school systems. As H. D. Meyer (2002b) depicted it:

> …the tide seems to have turned. More conventional control- and command-oriented managerial thinking (frequently originating in the world of private enterprise) seems to be back, welcomed under labels such as "new managerialism", or "entrepreneurial management". We notice a stronger emphasis on organizational effectiveness, accountability, capacity building, and standardization (p. 516)

Definition of new managerialism. New managerialism, according to Goldspink (2007), was simply "an application of business management principles to public institutions" (p. 27). The term was interchangeable with "new public management" (Bottery, 1996; Goldspink, 2007). To differentiate new managerialism with existing knowledge about public management, Gunter and Fitzgerald (2013) pointed out that the new public management was different in the sense that: a) "new" meant bringing newness to generate reforms, and to "signal a break with traditional cultures and professional practices"(p. 214); b) "public" meant introducing business methods and cultures into public service department; c) "management" emphasized that "on making bureaucracy work better, not least through shifts in power relations known as managerialism". (Ball, 1998) pointed out that new managerialism was "the insertion of the theories and techniques of business management and the `cult of excellence' into public sector institutions" (p. 123), and constantly stresses the importance of quality. Morley and Rassool (1999) noted that new managerialism "has been a generic transformational device designed to restructure and reorient public service provision" (p. 61).

There was no universal guidance for achieving new managerialism. As Gunter and Fitzgerald (2013) summarized, there was "no blueprint or rule book" (pp. 214-215), "it is

through accounts of practice that NPM can be identified, where scholarly debate can show the differences between techniques and shifts in power relations" (p. 214). Morley and Rassool (2000) stated that new managerialism "central to new managerialism is the promotion of a corporate mission, with goals, targets, monitoring procedures and performance measurement" (p.173). In literature, the practices of new managerialism were frequently related to quality control, performance, standardization, accountability, and data. Codd (2005) noted that "managerialism is preoccupied, if not obsessed, with the notion of 'quality'" (p. 200), and "within the discourse of managerialism, quality has been linked to accountability through a process that is called quality assurance" (p. 201). Ball (1997) noted that "professionality is replaced by accountability, collegiality by competition and interpersonal performative comparison" (p.261). Gunter and Fitzgerald (2013) stated that the new managerialism needs to be secured through performance regimes based on "beliefs in and about data" (p.214), and "the employment and training of new personnel who produce and interpret the statistics" (p.214). Hood (1991) offered seven doctrinal components to practically pin down new managerialism, which included:

* 'hands-on' professional management in the public sector;
* explicit standards and measures of performance;
* greater emphasis on output controls;
* the break-up of monolithic into smaller manageable units;
* shifts to greater competition in the public sector;
* the stress on private-sector styles of management practice;
* the stress on greater discipline and parsimony in resource use (pp. 4-5).

Hood's seven doctrinal components were frequently cited in new managerialism articles. As others have noted (e.g., Bottery, 1996), the seven components did not equally exhibit in practice of different public sectors. In the U.S. public education, two major new managerialism

initiatives—standards-based reform, and performance-based accountability movement—were promoted and implemented top-down by the policy makers at federal and state level.

In both practice and literature, the standards-based reform and performance-based accountability closely intertwine with each other. Here I tried to associate them with different component of Hood's seven doctrinal components. Standards-based reform, at least, related to Hood's (1991) second component of new managerialism, which, respectively, suggested "definition of goals, targets, indicators of success, preferably expressed in quantitative terms, especially for professional services" (p. 4). The performance-based accountability, however, associated with Hood's first, fifth and sixth components, which suggested "active, visible, discretionary control of organizations from named persons at the top, 'free to manage'" (p. 4), "move to term contracts and public tendering procedures" (p. 4), and "move away from military-style 'public service ethic', greater flexibility in hiring and rewards; greater use of PR techniques" (p. 5). Both standards-based reform and performance-based accountability movement have overhauled the dynamics of U.S. educational systems. Standards-based reform was designed to align educational elements and minimize variability; Performance-based accountability was implemented to assure the responsibility of each educator with accompanying consequences.

Standards-based Reform—Restructuring the Entire System

The widespread application of managerialism meant that educational institutions, like many other public service sections, were no longer exempted from external inspections and controls. Although the role of federal and state governments was becoming increasingly important, the goal of standards-based reform was not just about the government involvement. Instead, it was about an aligned system of standards and instructional guidance at all levels of the educational enterprise (Mason, 2003). In the literature, the "systemic reform/change" was

deemed as interchangeable with "standards-based reform", and as Knapp (1997) noted, the two terms were virtually synonymous. Standards-based reform was expected highly by both the policy makers and the public, could standards-based reform really change education systems?

The conception of standards-based reform. Sarason (1990) once argued: the proponents of educational reform do not talk about changing the educational system (p. 13). According to this judgment, recommendations of education remained the same for decades and did not target the whole system. Pointing out that past reforms of education were limited by their scope, Smith and O'Day (1990) suggested a "fragmented, complex, multi-layered educational policy system in which they (schools) are embedded (p. 237)" was a "fundamental barrier to developing and sustaining successful schools in the USA" (p. 237). They argued that the state was the one which can target the whole system, and through an alignment of intended changes on instruction, curriculum, tests, and professional development by states and districts, the reform efforts could become more consistent and effective, because the systemic change could overcome the looseness of the school system and penetrate all the way to the classroom level. Smith and O'Day's (1990) framework of standards-based reform was further articulated by other researchers. (Morley & Rassool, 2000) noted that "underpinning this move was the desire to standardize and reduce product variety in education. Quality can only be evaluated when like is compared with like" (p. 172). Willms (2000) provided examples of necessary evaluation techniques to monitor school performance in the era of standardization. Massell, Kirst, and Hoppe (1997) suggested that although standards-based reform could have various forms, there were three key features of standards-based reform:

> Challenging academic standards set by the state that would specify what all students should know and be able to do;
>
> Aligned policies, such as testing, accountability, teacher certification and professional development tied to the new, challenging standards; and

> Restructured educational governance to enable local teachers and schools to decide upon the specific instructional programs they would use to achieve the standards (p. 1).

Chatterji's (2002) definition of three defining components of systemic change coincided with Massell et al.'s (1997) summary of three key features. Swanson and Stevenson (2002) stressed standards-based reform had "a process-driven conception of educational change that explicitly links schooling inputs and policy drivers to student outcomes through clearly defined mechanisms" (p. 3). Clune (1998) articulated a continuous causal sequence of standards-based reform:

> Systemic reform (SR), through its purposeful activities, leads to
> Systemic policy (SP), which leads to
> A rigorous implemented curriculum (SC) for all students, which leads to
> Measured high student achievement (SA) in the curriculum as taught (p. 2)

According to Clune (1998), the systemic policy in the continuous causal sequence is the driving force for change, which covers "full range of influential policies in the area" (p. 11). Standards-based curricula is the "touchstone for policy alignment"(p. 9), which refers to "the material actually conveyed to students in classrooms and the instructional methods by which it is taught" (p. 12). And systemic student achievement should be measured with "gain on a student assessment in some way aligned with the reform" (p. 12).

Challenges to standards-based reform. Given the development of systemic change, researchers sought to understand how standards-based policy and practice played out at the district and school levels and influenced teaching and learning in the classroom (Mason, 2002; 2003). In the U.S., the ideas of standards-based reform were first introduced in the 1994 revision of Title I of the Elementary and Secondary Education Act. The revised statute required states to "prescribe standards in all subject areas which describe in detail the skills and curricular content that students are expected to learn (Willms, 2000, p. 238)" and "to measure students'

performance against these standards, and to assess students' overall mastery of the state's goals as advanced, proficient, or partially proficient" (Willms, 2000, p. 238).

D. Cohen (1995) argued that standards-based reformers were challenged to make more coherent guidance for instruction. Standardization was not simple, as D. Cohen stated, "there has been a broad movement toward intellectually more ambitious instruction, but there also has been great variability within that movement, as well as many contrary and simply different movements" (p. 12). At the state level, the new guidance for instruction delivered from states to district central offices was often mixed and conflicting. Correspondingly at the district level, local central offices picked the direction of reform simply according to their taste—some districts implemented some certain reform activities as they happened to be in accordance with their long-hold plans, otherwise, central offices would ignore the reform. Similar scenarios were repeated within districts, principals' responses to central offices were also varied significantly, some embraced the changes, some stayed with convention, and some took a neutral stance. D. Cohen proposed four explanations for the increased inconsistency under standards-based reform movement. First, state instructional policy did not constrain local instructional policy making. Second, the differences of implementation were amplified because the U.S. school system was too fragmented to promote "systemic" reform guidance—as reform guidance were "picked up by an astonishing variety of organizations-all concerned with schools but each concerned in its own way" (p. 13). Third, reform proposals themselves introduced chaos through pointing in too many different directions at once. In practice, some states advanced new guidance to strengthen the state regulation but at the same time, did not care to cancel the old guidance. Lastly, systemic reformers who promoted demanding standards had too much disagreement on purposes and

methods. Unsettled topics such as "what such standards are, how they might be developed, and how they might be used (p. 13)" brought great confusion to practitioners.

D. Cohen (1995) was not the only one who warned about the perils of standards-based reform, Schmoker and Marzano (1999) pointed out that "the perception of a common, coherent program of teaching and learning is a delusion" (p. 19). They noted that teachers made "independent and idiosyncratic (p. 19)" instructional decisions even when highly structured textbooks were used as the basis for a curriculum.

Elmore (2000) cautioned that standards-based reform could not resolve various issues in the school systems simply by "tinkering with the existing institutional structure" (p. 8). He argued that the logic of standards-based reform was "fundamentally at odds with the logic of loose-coupling, and this difference is not likely to be resolved in the usual way, by simply bending and assimilating the new policy into the existing institutional structure" (p. 8). Elmore noted three conflicts between logic of loose-coupling and the logic of standards-based reform:

> First and most surprisingly, standards-based reform violates the fundamental premise of loose coupling—buffering the technical core from interference by external forces.
>
> Second, standards-based reform hits at a critical weakness of the existing institutional structure, namely its inability to account for why certain students master academic content and can demonstrate academic performance while others do not.
>
> Third, standards-based reform undermines a basic premise of local governance of education because it identifies schools, not school districts, as the primary unit of accountability in virtually all state accountability systems (pp. 8-10).

The empirical finding from Shen and Ma's (2006) study seemed to partially supported D. Cohen and Elmore's points. They investigated "how the systemic change theory transpires when it was applied to the technical core of teaching and learning within a loosely coupled system (p. 235)" with national representative data. The findings suggested an increasingly attenuating association from the states' guidelines, to the schools' curriculum, and to classroom teachers'

instruction. In other words, standards-based reform was able to penetrate all the way to the school level, but stopped just at the classroom door.

Swanson and Stevenson (2002), however, had some different findings. They applied Hierarchical Linear Modeling (HLM) to examine the linkage between state policy and teacher's classroom practice under standards-based reform. Their findings suggested that state policy had a significant effect on classroom instruction with a standards-based model of mathematics education, albeit the small effect size.

Overall, although there were a few studies evidenced some positive changes of instructional practice (Bridge, Compton, Hall, & Cantrell, 1997; Burian-Fitzgerald, McGrath & Plisko, 2003; Swanson & Stevenson, 2002; Clarke, Shore, Rhodes, Abrams, Miao, & Li, 2003) and improved student achievement (Butty, 2001; Erickson & Niess, 1996; Hickey, Moore, & Pellegrino, 2001). Shen and Ma's (2006) conclusion was confirmed by more studies, which all suggested that standards-based policy or curricula had limited influence on teachers' instructional practice (see Falk & Drayton; 2004; S. G. Grant, 2001; T. J. Grant & Kline, 2000; Hamilton, McCaffrey, Stecher, Klein, Robyn & Bugliari, 2003; Spillane & Jennings, 1997, etc). Moreover, low coherence between instructional guidance and teacher's in-class practice was not the only issue of standards-based reform. A few studies reported that the alignment of state standards with state assessments was weak as well (Polikoff, Porter, & Smithson, 2011; Webb, 2005, 2006).

To sum up, in the past decades, while the conception of standards-based reform seemed promising, the research on standards-based reform suggested mixed findings. There was no clear evidence to support the effectiveness of standards-based reform.

Accountability Movement—a Multifaceted Conundrum

Standards-based reform had great potential to bring better education to the public. Theoretically speaking, the technical core of school—the ambiguity in instructional or curriculum related decision making—could be transcended by alignments of instructional guidance, curriculum development, standardized testing, and professional development. However, in reality, the influence of standards-based reform on districts, schools, and teachers was still attenuated or buffered (D. Cohen, 1995; Schmoker & Marzano, 1999). It is noteworthy that a body of studies indicated that: standards-based reform was able to penetrate to the school gates, but not the classroom doors (e.g., Chance & Anderson, 2003; Elmore, 2000). Therefore, it was suggested that more powerful tools were needed for systemic change to influence classroom-level activities (Shen & Ma, 2006).

Background of accountability movement. Accountability was an idea originated from industrial models. Based on a "modernist, rationalist (Morley & Rassool, 2000,p. 172)" belief that "the complexities of the social world can be measured and recorded with the appropriate instruments and technologies" (p.172), performance-based accountability was widely introduced to public service provision. In the field of education, accountability had become an especially fundamental part of education policy. As Harris (2007) noted, "The idea of accountability for performance has a firm grip on education policy in virtually every industrialised democracy" (p.41).

Successful implementations of standards-based reform could not be maintained without including performance-based accountability as a part of the "systemic policy" (Clune, 1998). Mason (2003) stated that the core idea of standards-based reform was an aligned system of standards and instructional guidance at all levels of the educational enterprise, reinforced by

accountability measures based on mandatory standardized testing. In the era of accountability, the attribution of "high-stakes" received increasing emphasis in standardized tests (Supovitz,2009). According to a review of half-century of school reform by the president of Center on Education Policy (Jennings, 2012a), the focus of the current standards-based reform had switched from standards to performance. As he noted:

> The original purpose of the standards-based reform movement was to identify what students should know and be able to do at specific grade levels and to measure whether they were mastering that content. As the movement matured, it took on the additional purpose of applying consequences to schools whose students did not show mastery. In this way the standards movement morphed into test-driven accountability (p. 5).

Holding schools and districts accountable for testing is not a new idea, as even before the enactment of NCLB in 2002. Many states already had their own accountability systems. Between 1995 to 2000, 11 states had strong accountability systems, 15 states had moderate accountability systems, and 4 states had weak accountability systems (see Lee & Wong, 2004). However, NCLB unprecedentedly and dramatically increased the federal power in U.S. education. It was the first time in U.S. history, that each state was required to set rigid "challenging" standards for all students in reading and math, which were expected to bring students to meet the proficient level in 2013-14 school year. Another big change brought up by NCLB was that states were required to attach consequences to performance, which was missing in many former state accountability systems (Hanushek & Raymond, 2005). Under NCLB, schools are held accountable for Adequate Yearly Progress (AYP) via standardized tests. If a school continually fails to meet AYP, it would face a series of sanctions, from offering students choices to other schools to shutting down of schools.

Although NCLB attempted to drive changes at the national level, state and local authorities were still playing important roles in many decision-making areas. The stringency level of NCLB accountability systems varies by state because of the different combinations of the states' decisions regarding their status quo, as some states have the flexibility to develop their own accountability schemes (Hanushek & Raymond, 2005; McCombs, Kirby, Barney, Darilek, & Mageest, 2004; Wei, 2012a). As Wei (2012a) noted, the federal law "allows states to design some parts of their accountability plans to adapt to local circumstances" (p. 269). Loeb and Strunk (2007) recommended states give local authorities more flexibility in decision-making areas such as revenue raising and teacher hiring to maximum the benefit of accountability policy.

Changes Made by Accountability Policy. Dee, Jacob, Hoxby, and Ladd (2010) stated: "NCLB has been extremely controversial from its inception" (p. 150). The proponents of NCLB believe that accountability is the key to bring more effective schools in the existing educational system. The opponents, however, argue that NCLB fails to make real changes in student performance. Instead, it creates more inequality and more confusion to conventional values in education.

Accountability and student achievement. Whether accountability really improved the student achievement is the most important question. Existing studies support the proposal that the accountability and high-stake testing did bring higher student performance.

Before NCLB, researchers evaluated the effect of accountability on student performance by grouping states into different categories—such as states which did not have accountability, states which had different types of accountability, and states which have varied extent of sanctions attached student performance. Overall, overwhelming evidence showed that after control for various background variables, a stringency level of state accountability system was

associated with a higher student pre-NCLB performance gain (Braun, 2004; Rosenshine, 2003; Carnoy & Loeb, 2002; Hanushek & Raymond, 2005; Wong et al., 2009). Only a few questioned the real influence of accountability on student achievement (Amrein & Berliner, 2002; 2003).

In December 2006, the U.S. Department of Education (2006) commented that "because of No Child Left Behind's accountability provisions, schools and parents are getting the information and help they need to focus attention and resources on the children who need it most—and it's working". The statement that "accountability was working" was supported by a body of studies, albeit some pointed out that the reliability of state testing and accompanying proficiency standards themselves were questionable (see Rothstein, 2008; Stecher & Barron, 1999). Researchers evaluated NCLB accountability by looking at the change of direction of test score trend before and after NCLB, because without an identification of trend changes, it was hard to tell "whether gains in achievement after NCLB were just a continuation of previous trends" (N. Chudowsky, V. Chudowsky, & Kober, 2007, p. 12). Center on Education Policy (CEP), claimed themselves to have conducted the "most comprehensive study of trends in state test scores since NCLB" (N. Chudowsky et al., 2007, p. 9) at the time, studied the pre-and post-NCLB trends of student state tests in all 50 states. Using percentage proficient as one measure of achievement and a statistical tool called effect size as a second measure, CEP was able to capture variations below and above the proficient level. They found that after the enactment of NCLB, student's achievement levels had been raising continuously in most states. Furthermore, the improvement of student achievement gain began to accelerate after NCLB. Dee and Jacob (2011) applied a statistical tool called comparative interrupted time series (CITS) to investigate the impact of NCLB, by comparing the 1992-2008 trends of student performance in states which had accountability before NCLB and in states who did not have accountability. Dee and Jacob

reported that NCLB statistically improved average math performance in both 4th grade and 8th grade, especially among 8th graders from traditionally low-achieving group received largest gain. Similar patterns of positive NCLB effect was confirmed in many other studies (Ballou & Springer, 2008; Lauen & Gaddis, 2012; Reback, Rockoff, & Schwartz, 2011).

Accountability and achievement gap. Although the accountability system, as reported by the most studies, had positive impact on overall student achievement, it has been argued that the accountability was not effective in narrowing achievement gap between subgroups. Existing studies showed mixed findings, albeit more studies suggested that accountability had no significant influence. Hanushek and Raymond (2005) reported that under pre-NCLB state accountability system, white students had higher gains relative to black and Hispanic students on each of the tests, Hispanic students proportionally gained most in the National Assessment of Educational Progress (NAEP) performance, while black students gained least. Hanushek and Raymond cautioned that the pre-NCLB accountability systems, although they improved overall performance, had widen the racial achievement gap otherwise, especially the gap between black and white students. They recommended the newer accountability under NCLB employed more additional policies and resources to address "more equal outcomes across groups" (p. 32). In the same vein, Lee and Wong (2004) noted that state pre-NCLB accountability failed to narrow the achievement gaps among racial and socioeconomic groups. Similar to Hanushek and Raymond's (2005) findings, Lee and Wong noted that Hispanic students in strict accountability states had improved performance. However, the improvement disappeared after the socioeconomic status was controlled. Lee and Wong therefore, argued that more resources need to be assigned to achieve equity among racial and socioeconomic groups. There was still some evidence to support the effect of pre-NCLB accountably on racial equality. For example, Carnoy and Loeb (2002)

found Blacks and Hispanics, between 1996 and 2000, had larger growth than Whites under accountability policy, especially for the 8th graders.

NCLB pressured states not only to bring better overall performance, but also to eliminate achievement disparities in reading and mathematics among subgroups. N. Chudowsky et al. (2007) found that among 38 states with necessary Post-NCLB state test data, 14 states had narrowed Black-White reading gaps, 12 states had narrowed Black-White math gaps; 13 states had narrowed Hispanic-White reading gap, and 11 had had narrowed Hispanic-White math gap. For students with low-income, the reading gaps were narrowed in 15 states among 31 states with necessary data , while math gaps were narrowed in 13 states among 29 states. N. Chudowsky et al. (2007) noted that gaps seemed to be narrowed in most states, and few states had widened gaps after NCLB, the achievement gaps among different subgroups were still sizable, which require long-term attention to eliminate them. Similarly, Dee and Jacob (2011) found that NCLB increased achievement of both black and Hispanic students significantly in math. Dee and Jacob proposed a plausible explanation for the variation of the effectiveness of NCLB among states. They suggested that NCLB was more effective on disadvantage students in "states with greater number of such children" (p. 172). In North Carolina, Lauen and Gaddis, (2012) reported that NCLB had small positive effects on Black and poor students, but larger positive effect on Hispanics between 2000 and 2008. However, another study conducted in Florida by Figlio, Rouse, and Schlosser (2009), found NCLB did not have significant effect on the performance of Black, Hispanic, and low SES students in Florida.

Students with disabilities and limited-English-proficient students (note: there were very few studies that had provided solid evidence of NCLB effects on this group) used to be exempted from state standard tests or given a different tests before NCLB (N. Chudowsky et al., 2007;

National Research Council, 1997). Under NCLB, all students, including students with disabilities were expected to reach proficiency with same high standards (N. Chudowsky et al., 2007; Wei, 2012b). Wei (2012b) found that NCLB did not have significant influence on California Standards Tests Scores (CST) performance of students with disabilities.

Accountability and school dynamics. Lee and Wong (2004) found that strict accountability did not change the distributions of school expenditures, class size, and qualified teachers in 90s. However, it seemed that NCLB fundamentally changed the dynamics of school work. Firestone (2009) observed that under the accountability movement, districts and schools gradually shifted from loosely coupled culture to accountability culture. During the 20th century, the educational goals in the U.S. tended to swing between academic and vocational, and between social and personal goals (Cuban, 1990; Goodlad, 1994; Shen, 1999). The balance seemed to be upset by the deepening of accountability movement. Emphasis on academic goals has become a predominant phenomenon since the 1990s (Shen, Palmer, & Crawford, 2005). Xia, Gao, and Shen (2015) studied principals' goals from 1990 to 2007 with a national representative data. They found that academic goals had been consistently rated as the most important educational goals to principals, while personal goals had witnessed a precipitous drop. It was noted principals were becoming invisible managers and more disconnected from teachers after NCLB (Hargreaves & Goodson, 2006).

Accountability movement also resulted in palpable changes in teachers' instructional behavior and allocation of instructional time. "Teaching to the test (Firestone, Mayrowetz, & Fairman, 1998)" began to prevail classrooms. It was found that NCLB shifted the allocation of instructional time toward math and reading (Dee & Jacob, 2011); teachers spent less time on social studies in those states that did not mandate testing on social studies (Fitchett et al., 2014);

and teachers spent more efforts on "bubble kids"—those who were just about to meet the proficiency—but much less efforts on kids further below the threshold (Booher-Jennings, 2005; Bracey, 2008; Neal & Schanzenbach, 2007). Desimone (2013) noted "the target of instruction has shifted from standards themselves to standardized test results, often exclusively" (p. 61). Teacher's "trade-off" schemes brought some worrisome consequences on students. It was noted that students began to "lose intrinsic interest in subject matter, learn at only a superficial level, and fail to develop a desire for future learning" (Sheldon & Biddle, 1998, p. 1). The development of high-ability students in low-achieving schools was suppressed by NCLB (Krieg, 2008).

Grissom et al. (2014) found that, according to teachers' perceptions, NCLB increased teachers' classroom control and administrator support to teachers, however, NCLB reduced teacher cooperation. It was also found that NCLB had no impact on teacher job satisfaction and job commitment. Sheldon and Biddle (1998) warned about the introduction of performance-based sanctioning systems. They argued that teachers could act "controlling, unresponsive to individual students, and alienated (p. 1)", they added, "they are likely to experience resentment and loss of morale, to engage in superficial conformity, and (eventually) to quit their jobs as teachers" (p. 6).

To sum up, although sufficient literature suggested that accountability could improve student performance, there were still a lot of worrisome doubt about its negative impact on particular student groups, teachers, and principals. Some even argued that standards-based reform before NCLB, with less accountability, was more successful. Desimone (2013) interviewed 60 teachers, 32 principals, 14 district administrators, and 7 state officials. For conclusion, she put it:

> Earlier attempts at standards based reform and accountability … were more closely aligned with the theoretical vision of standards-based reform than were later

> manifestations as codified under NCLB, which has moved us away from school and teacher discretion, and placed less emphasis on the need for teacher buy-in. Instead, rewards and sanctions have become the primary mechanism for fostering change (p. 61).

Rise of Data-informed Decision-making

There is little doubt that data played a vital role in U.S. Education. Melissa Roderick, a professor at the University of Chicago and a senior director of the UChicago Consortium on School Research put it: "…not charters, or small schools, or standards, or accountability, or any of the other reform platforms that have shaped American education over the last several decades. It is, simply, data" (Roderick, 2012, p. 1).

Roderick's statement is thought-provoking. It seemed that the rise of data led unwittingly to a new wave of education reform. Both the standards-based reform and the performance-based accountability movement put great emphasize on data. New Managerialist deemed data as the centerpiece of management. Data is believed to bring profound changes to the structures of educational organizations and their management (Gunter & Fitzgerald, 2013; Morley & Rassool, 2000, etc). Datnow and Hubbard (2015) argued that though a process of continuous improvement, data could "become infused into the structure and culture of the organization" (p.8). Successful managerialism required production of data, interpretation of data, and most importantly, beliefs in data (Datnow & Hubbard, 2015; Gunter & Fitzgerald, 2013).

NCLB highlighted the responsibility of states, districts, and schools to use data for school evaluation and improvement (Mandinach, Honey & Light, 2006). Following NCLB, the Race to the Top fund program in 2009 requires states receiving RttT grants to improve teacher and principal effectiveness based on student performance data. States were required to "conduct annual evaluations of teachers and principals that include timely and constructive feedback; as part of such evaluations, provide teachers and principals with data on student growth for their

students, classes, and schools" (U.S. Department of Education, 2009, p. 9). Both NCLB and RttT stressed the data as the essential tools in education reform. Mountains of data were generated from standard testing and accompanying accountability measures (Celio & Harvey, 2005). With the accumulation of data, it became possible for educators to operationalize reform concepts into educational practice in reality (Thornton & Perreault, 2002).

Data-informed decision-making, and data-driven decision making are two phrases mostly cited to describe the activities that educators improve education quality with data. However, some have noted nuances between the two in the literature. Shen and Cooley (2008) recommended that data-informed decision-making was more appropriate, for the reason that other than data, rational, political, and moral inputs were also important to decision making. Therefore, it was preferred that data *inform* rather than *drive* decision making.

I was aware of the subtle differences between data-informed decision-making and data-driven decision making during this literature review, and chose to use data-informed decision-making in the following writings. Nevertheless, the two concepts were deemed interchangeable for the reason of gathering as many studies as possible. Overall, the nuance between data-informed decision-making and data-driven decision making did not affect general understanding of the concepts in the literature.

Decision Making in Previous Times

Decision making in the field of education could be vague in previous times. A classic study conducted by M. Cohen and colleagues (1972) in 70s demonstrated remarkable ambiguity in a decision making process in an educational organization. M. Cohen and colleagues used the metaphor garbage can to present how irrationally and randomly that decision making were made in a university. They proved that the elements of decision making were disconnected through a

garbage can process of decision making, which failed to offer random rather than good decision for solving issues. Similar patterns were pervasive in schools, sufficient literature had argued that evaluation and inspection worked more as ritual than as control tools in schools, and productivity of teachers rarely received proper measurement (see Cuban, 1990; J. W. Meyer & Rowan, 1978).

At the present time, perhaps no one will doubt how much educational systems relied on student academic performance to evaluate schools, principals, and teachers. Back in the 70s, only one of 34 superintendents claimed to use achievement data to evaluate schools, according to a survey conducted in the San Francisco Bay Area (J. W. Meyer & Rowan, 1978). But in 2003, 63% of superintendents believed "raising student achievement is the biggest part of principals' performance evaluations," according to a national-level report sponsored by the Wallace Foundation (Farkas, Johnson, & Duffett, 2003). Despite the differences in samples and regarding questions between two surveys, the remarkable disparity in educators' preference on data over the past decades implied a drastic change in education—data is becoming increasingly important in school decision making. Although most states were still incapable of collecting, analyzing, and reporting data on individual schools two decades ago (Elmore, 2000), thanks to the modern technology and standard based reform and accountability movement, data is now widely applied to reduce ambiguity of decision making across education systems. More importantly, data are able to connecting educational practitioners at classroom level, school level, and district level.

Using Data for Continuous Educational Improvement

The rise of data-informed decision-making implies "a fundamental philosophical shift from data for compliance to the principles of data for continuous improvement" (Mandinach, 2012, p. 72), while accountability and compliance are no longer the most important goal of using data (Daly, 2009; Jimerson & McGhee, 2013). Using data to make informed school decisions

revealed a new perspective to resolve issues in education—continuous improvement (see Anderson & Kumari, 2009; Fullan, 2007; Halverson, Grigg, Prichett, & Thomas, 2007; Sutherland, 2004).

Compared to the previous top-down reform endeavors that were initiated by external stakeholders, continuous improvement stressed engaging schools in endogenous dynamic which was believed to fundamentally change schools into learning organizations, where practitioners develop "shared understanding and commitment to achieve high level outcomes for all students" (Fullan, 2007, p5). Using data for continuous educational improvement is inspired by successful models from industry and manufacturing sectors. There have been sufficient in-depth discussions about applying industry-based concepts such as Continuous Improvement, Total Quality Control and Kaizen to school reform in the literature (see Bernhardt, 2013; Marsh, Pane, & Hamilton, 2006; Ikemoto & Marsh, 2007; Morley & Rassool, 2000). Based on the traditional, four-step Continuous Improvement Model developed by Deming (1986)—plan-do-check/study-act/adjust, Bernhardt (2013) suggested a plan-implement-evaluate-improve strategy for continuous school improvement.

Using data for decision-making is not a new idea to school teachers. In the past decades, improving the quality of instruction with student performance data has been widely introduced to classroom (Shen & Cooley, 2008; Ikemoto & Marsh, 2007). Nevertheless, the concept is far beyond simple data utilization at classroom level. Following the standards-based reform and performance-based accountability, data is more and more accepted as a tool to change the entire education system. It is expected that teachers, principals, and central offices work cooperatively and systematically to use data to inform a range of decisions for improvement (Ikemoto & Marsh, 2007). The complexity is that it requires more types of data input sources, more advanced

data analysis schemes, and most importantly, more combined and systemic effort from each level of educators in the school systems.

Existing literature suggested that data-informed decision-making is predominately used for continuous improvement of curriculum and instructional practices. According to a report sponsored by the U.S. Department of Education, most common school-level uses of data described by teachers and school leaders were "school improvement planning, curriculum decisions, and placement or grouping of students for instruction or support services" (Means, et al., 2010, p. xiv). Likewise, based on a survey of 68 schools, Supovitz and Klein (2003) summarized three common uses of data: a) informing instruction (including identifying lesson objectives, focusing instruction, and aligning lessons with standards); b) developing assistance plans for low-performing students (including identifying students and monitoring student progress); c) setting goals and targets (including setting ambitious annual improvement goals as well as finer-grained goals). Firestone and González (2007) concluded that schools or districts using data for "guidance for action"—which meant using data to monitor content and delivery of instruction, performance of students, and quality of curriculum; "enlightenment"—referred to "situations where information orients users to educational issues or ideas" (p. 137); and "mobilize support"—referred to school and district leaders use of data to "obtain support from teachers or the public for a specific change" (p. 138).

Data-informed decision-making may not only support improvement at district and school level, but it may also promote a teacher's professional development. Gummer and Mandinach (2015) used *data literacy for teaching* to evaluate teacher capability to "transform information into actionable instructional knowledge and practices by collecting, analyzing, and interpreting all types of data to help determine instructional steps" (p.2). Pardini (2000) noted that teachers

were more likely to modify their instructional strategies when they had current information about the skill levels and proficiencies of their students. Datnow, Park, and Wohlstetter (2007) found that when data suggested students need additional supports, teachers would adjust their teaching strategies accordingly. Massell (2001) found that after teachers received enough exposure to data, their view toward capabilities of some underperforming groups improved. More studies suggested the power of teacher collaboration or professional learning networking in strengthening teacher capacity for data-informed decision-making (Cosner, 2011; Datnow, Park, Kennedy-Lewis, 2013).

No doubt data-informed decision-making, in conception and in many practices, can be a key to continuous improvement at all levels of educational systems, yet it was discovered that data using, in many scenarios of the real world, is still "superficial and focused more on accountability and satisfying the state's requirements than a systematic investigation into conditions that support and detract from the teaching and learning process" (Shen & Cooley, 2008, p. 320). If it is expected the circle of continuous improvement will be implemented across all levels of education systems, maintaining a balance between the accountability and improvement functions of data-informed decision-making become critical (Jimerson, 2014). Teachers especially preferred to use data for the purpose of classroom improvement, rather than for the purpose of accountability, or compliance (Daly, 2009). To transform existing accountability system into a system facilitating data use at teacher level, Jennings (2012b) advocated focusing on five areas of accountability systems, including: a) varying amounts of pressure; b) the locus of pressure varies across systems; c) distributional goals; d) varying features of assessments; and e) applying different system according to scope (see pp. 5-17).

Data-informed Decision-making in Loosely-coupled System

The recommendation for educational data-informed decision-making sounds promising, but is too vague to practice. Especially, determining how data and local decision making are related remains unexplained in policy text (Spillane, 2012). Shen and Cooley (2008) pointed out that data-informed decision-making alone would not renew the educational system, if a) not enough coherent changes are made to ensure data-informed decision-making knowledge and skills of stakeholders; and b) missing the corresponding structural arrangement to facilitate and harness the power of data-informed decision-making. Data is believed to fundamentally restructure the education system. However, barriers to systemic and consistent data utilization root deeply in the education system.

It is imprudent for schools and districts to continue to ignore the importance of data-informed decision-making. Teachers used to complain about "reform *du jour* (Thornton & Perreault, 2002, p. 86)" when new programs were "implemented, modified, and then replaced or continued without appropriate, data-based evaluation and decision making" (p. 86). A good deal of literature documented how districts could shape data-informed decision-making at school and classroom level (e.g., Anderson et al., 2010; Togneri & Anderson, 2003). However, more literature suggested that it was challenging for stakeholders to work collaboratively with data in a loosely coupled school setting. Evidence from previous studies showed that teachers, schools and districts tended to collect data differently (Fuchs, 2004; Honig & Coburn, 2008; Shen, et al, 2010); interpret data differently (Ingram, Louis, and Schroeder, 2004; Lachat, Williams & Smith, 2006; Rudner & Boston, 2003); and utilize data differently (Diamond & Cooper, 2007; Hess & Fullerton, 2009; Ikemoto & Marsh, 2007; Little, 2012).

Literature has argued that the incoherent data-informed decision-making may be a result of different mental models (Datnow & Hubbard, 2015; Jimerson, 2014). It is found that educators have their own preference of data sources and data applications. For example, Shen and colleagues (2010) interviewed 16 principals in four urban school districts in Michigan to investigate what data principals used and how they used data. They found that principals used student achievement data more frequently than school process data and community/student background data. The decisions principals made based on the data were predominantly for accountability rather than for learning. In other words, principals tended to use data *of* learning, rather than data *for* learning. However, they did confirm that districts had great influence on promoting data-informed decision-making for curriculum- and instruction-related issues among principals. Thorn (2001) discussed distinctions between district-level and school-level data-informed decision-making, suggesting that central offices tended to use summative data to evaluate program, while schools prefer formative evaluation to improve instruction.

Coburn and Talbert (2006) studied the differences among individuals' conceptions of valid evidence. They suggested that the use of evidence was attributed to "individuals' work in different parts of the organization and their involvement in different facets of the district's reform history" (p. 470). They observed that district administrators were interested in student performance trajectories on standardized tests for policy assessments and accountability measurement, while teachers tend to only look at student performance in discipline content to evaluate their instructions. What's more, Coburn and Talbert believed that the reform efforts from higher levels could be constrained by local consensus, ideologies, or default working responsibility that created by previous reform movements.

Young (2006) attributed the disparity of data-informed decision-making at different levels to their misalignment in agenda-setting. Agenda-setting, according to Young, was "leadership in articulating rationale, setting expectations, and structuring time and teachers' learning about data" (p. 523). In his qualitative study to explore the influence of school leadership on teachers' data uses, Young noted that "agenda setting that engages the non-rational aspects of the system: norms, values, and capacity situated in role definitions and hierarchy (p. 545)" was a foundation for a rational system of data-informed instruction. Young suggested that school leaders could greatly mediate guidance or policies through agenda-setting. Moreover, school leaders' commitment and their skills underscore districts seeking to institute data-informed practices. Datnow, Park, and Wohlstetter (2007) attached great importance of explicit expectations and norms to data-informed decision-making.

Some studies argued that the inconsistency of data-informed decision-making stemmed from practitioners' incapability to collect, analyze, and interpret data. Dennis (2008) cited a case in which a teacher was unable to provide a program more conductive to the student's needs, because the teacher had not received training necessary to understand results of state standard test. Mason (2002) suggested that the willingness and capacity to use data was critical to teachers. Reeves and Burt (2006) suggested that while "teachers possess even less understanding of and appreciation for using data in decisions about classroom and school processes" (p.70), principals need targeted professional training on data collection, analysis, and interpretation. Thornton and Perreault (2002) pointed out that teachers' ability to use data was influenced by principals' understanding of data-informed decision-making. McNamara (2000) argued using school-related, real-world data to train new principals.

Central offices had the responsibility to promote data-informed decision-making and strengthen the capability of practitioners in using data in districts (Anderson et al., 2010; Leithwood, Louis, & Anderson, 2004). Firestone (2009) described a typical district with incoherent data support:

> Typically, teachers see average test scores, but often only months after students took the tests. Teachers are exhorted to do better, but there is no effort to break down scores to identify students or subject areas with correctable weaknesses (p. 672).

Means et al., (2009) highlighted the two most common ways that districts use to encourage school data-informed decision-making: a) incorporate data-informed decision-making practice into school improvement planning; and b) providing technical experts in systems, networks or databases who can help school staff get access to data from electronic systems (pp.14 -15). Datnow, Park, and Wohlstetter (2007) suggested designating data assistance managers to support system-wide data analysis at both district and school levels. They also suggested that it would be better if data mangers were "experienced, respected educators (p. 72)" instead of "technologists or statisticians" (p. 72).

Data generated from high-stake testing was questioned by practitioners for their instructional usefulness (Young, 2006; Stecher & Barron, 1999; Supovitz, 2012), stressfulness (Means, 2010), and unsynchronized feedback (Leithwood et al., 2004). Leithwood et al. (2004), Leithwood and Levin (2005), and Jennings (2012b) suggested districts take the responsibility to implement more comprehensive, multi-measure accountability systems to inform decision making at the classroom, school and district levels. In the same vein, Jimerson (2014) advocated focus on improvement functions of data-informed decision-making rather than accountability and monitoring.

Honig and colleagues published a series of papers regarding the incoherence of data using and conflicts of implementation of reform between schools and districts (Honig, 2003; Honig & Hatch, 2004; Honig, 2006; Coburn, Honig, & Stein, 2009; Honig, 2012). They advocated alleviating the policy incoherence between districts and schools by appropriate negotiation. On the one hand, they suggested school leaders "act as judges or informed, grounded interpreters of their multiple demands and school district central office administrators become interpreters and supporters of schools' local decisions" (Honig & Hatch, 2004, p. 19). On the other hand, Honig suggested administrators in central office operate as "boundary spanners" (Honig, 2006) and "principal coach" (Honig, 2012). To be "boundary spanners", district administrators should "work among the central office and schools and their community partners to broker new support relationships and otherwise enable implementation" (p. 357), and become "principal coaches", instead of monitoring the performance of school principals. District administrators should work closely with school principals to help them improve their practice.

From Loosely Coupled Organizations to Organizational Learning

The loose coupling theory suggests that educational organizations are loosely coupled by nature. Loose coupling keeps organizations stable and durable by protecting them from external changes. Over the past decades, loose coupling of educational organizations is deemed the major reason why the educational reform always failed.

Management practice in the public service sections has been profoundly molded by the new managerialism. Essentially, the education reform proposition of new managerialism tightens the systems by establishing standards and accountability. This kind of approach is especially powerful in reducing the misalignment and ambiguity in education, and exerting sufficient extrinsic pressure to stakeholders. Most important studies on student performance agreed that the

overall student performance were improved after NCLB. However, nothing comes without consequence. Because the new managerialism relies heavily on measurable indicators, bureaucratic control, and top-down external forces, it has great limitation in achieving sustainable and endogenous educational change. O'Day (2002) cautioned that the outcome-based, bureaucratic accountability led educators to focus less on information relevant to teaching and learning than on accountability requirement. It was shocking to find the systems were only superficially "tightened"—principals were becoming an even more "invisible" managers and had less communication with teachers (Hargreaves & Goodson, 2006); teachers were losing their morale in teaching, and engaged in superficial conformity (Sheldon & Biddle, 1998); and the top reason for using data was to satisfy accountability requirements (Shen & Cooley, 2008).

A Paradigmatic Debate on Tight Coupling and Loose Coupling

There is essentially a paradigmatic debate on how to take into account the relative advantage and disadvantage of loose coupling.

On one hand, the proponents of the managerialism paradigm argue that loose coupling was a problem and the key to educational improvement was to find mechanisms to tighten the educational system (Morley & Rassool, 2000). Both standards-based reform and performance-based accountability reform in the U.S. reflect the effort to tighten the loosely coupled system by publishing curriculum standards, developing accountability tests, and providing rewards and sanctions based on the results of the accountability test (Fusarelli, 2002; Smith & O'Day, 1990). Those who saw the relative disadvantage of loose coupling tended to advocate a top-down approach with a change model for the bureaucratic realm.

On the other hand, there are also those who believe the educational system was loosely coupled by nature and the key to educational improvement is to find ways to take advantage of

the loose coupling. In review of lessons from policy implementation, McLaughlin (1987) highlighted the importance of the will and capacity of educators to make changes in the success of reform. A bottom-up approach with a change model could take advantage of loose coupling nature of the system. For example, Goldspink (2007) proposed a model of school improvement that worked with, rather than against the loosely coupled system, which advocated a focus on "people, relationships and learning rather than structures and centrally determined standards for conformance" (p.27).

Table 1

A Comparison of the Managerialism Approach and the Organizational Learning Approach on Loose Coupling

	Loosely coupling	Managerialism		Organizational Learning
		Standards-based reform (since 1994)	Performance-based accountability (since 2002)	Continuous improvement
Basic Function	Buffering reform for localization (Weick, 1976; J. W. Meyer & Rowan, 1978); exempted educators from external inspections and controls (J. W. Meyer & Rowan, 1978); ensure the stability and durability of the system (Weick, 1976)	Restructuring—alignment of educational resources (Smith & O'Day, 1990)	Restructuring—hold educators accountable for student performance (Hanushek & Raymond, 2005)	Continuous and systemic improvement (Louis, 2007; Hord, 1997); teamwork for building professional capacity (Senge, 1990); ensuring flexibility and adaptability of improvement (Senge, 1990)
Motivation	Depends on personal preference because of volunteerism effect (Elmore, 2000)	Demoralized teachers, teachers acted superficial conformity (Sheldon & Biddle, 1998)		Self-motivated, committed to lifelong learning and students (Senge, 1990; Fullan, 2007); job satisfaction (Luthans, 2010)
Data	Rarely used (Cuban, 1990; J. W. Meyer & Rowan, 1978)	Mountains of data from testing (Celio & Harvey, 2005); focused more on data for accountability (Shen & Cooley, 2008); based on high-staking tests (Supovitz,2009)		"infused into the structure and culture of the organization" (Datnow and Hubbard, 2015)
Decision making	Irrational (M. Cohen, 1972); a lot of ambiguity (Firestone & Wilson, 1985); majorly for local needs (J. W. Meyer & Rowan, 1976)	For external standards (Desimone, 2013)	Reinforce traditions (Hargreaves & Shaw, 2000; Little, 1999); exclusively for standardized test (Desimone, 2013)	Create new knowledge and share ideas (Hubbard, Mehan, & Stein, 2006; Louis, 2007; Stoll et al., 2006); Focus on local needs (Louis, 2007); reinvent practice (Little, 1999)
Channel of control	Cultural linkage (Firestone & Wilson, 1985)	Bureaucratic control (Gunter & Fitzgerald, 2013)		Shared norms (Stoll et al., 2006); collegial and cultural linkages (Louis, 2007; Firestone, 2009; Luthans, 2010)
Accountability	Not accountable, because of the logic of confidence (J. W. Meyer & Rowan, 1978)	Some states started to have their own accountability policies (Lee & Wong, 2004)	Consequences are attached to performance in all states (Hanushek & Raymond, 2005; Jennings, 2012a).	Educators are hold accountable by themselves and for themselves (Dufour, 2004)

Table 1—Continued

	Loosely coupling	Managerialism Approach		Organizational Learning Approach
		Standards-based reform (since 1994)	Performance-based accountability (since 2002)	Continuous improvement
Principal/teacher relationship	A logic of confidence (J. W. Meyer & Rowan, 1978)	Disconnected and invisible manager (Hargreaves & Goodson, 2006); monitor compliance (Firestone & González, 2007)		Leaders are designers, coaches and stewards (Senge, 1990); teachers are reformers (Phillips, 2003)
Teacher Collaborations	Isolated (Goodlad & Klein, 1975; O'Day, 2002); experience of teachers tended to be private instead of shared (Lortie, 1975)	Reduced teacher cooperation (Grissom et al., 2014); alienated (Sheldon & Biddle, 1998);		Mutually supportive, interconnected (Senge, 1990); Team learning (Senge, 1990); interdependence is important to collaboration (Stoll et al., 2006); professional autonomy (Stoll et al., 2006)
Technical Core	Impossible to change (D. Cohen, 1990; Cuban, 1990; J. W. Meyer & Rowan, 1978)	Reform stopped at the classroom doors (D. Cohen, 1995; Elmore, 2000; Shen & Ma, 2006)	Superficially conformity (Sheldon & Biddle, 1998); Teaching to the test (Firestone, Mayrowetz, & Fairman, 1998)"	A consistent focus on teaching and learning through collaborative work, use data to evaluate progress over time (Stoll et al., 2006)

The development of Table 1 was based on the literature to compare the managerialism approach and organizational learning approach in several important areas. Obviously, both the managerialism approach and the organizational learning approach are designed to improve various issues of loose coupling—both approaches advocate changes; a focus on learning; and utilization of data, etc. However, compared to the more ambitious managerialism approach that aims to "restructure" the entire system, the organizational learning approach emphasized minor but sustainable changes initiated from inside of the system, through the collaboration of educators. Teachers are learners; learning facilitators; and reformers, they are self-motivated, accountable to themselves, and committed to learning. They are actively engaged in learning rather than superficial compliance, and obtained increased the professional capacity and job satisfaction through learning. Moreover, the organizational learning approach aims to improve, rather than eliminate some positive features of loose coupling, such as stability, emphasis on local needs, professional autonomy, diversity, and independence of teachers.

Data and Decision-making Driven by Different Paradigms

Data is at the centerpiece of continuous improvement (Datnow & Hubbard, 2015), as continuous improvement is the core of organizational learning. Different from the loosely coupled education organizations from last century which rarely used data, and Managerialism approach which relied mainly on mountains of data from standard testing, Organizational learning is all about data for learning. As Datnow and Hubbard (2015) described it, data use "infused into the structure and culture of the organization" (p. 8). As a reformer himself/herself, organizational leaners not only use data from standard tests, but also produce and collect data from other sources, such as parent surveys, school process data,

While the managerialism approach may strengthen the organizational decision makings by attaching more rationality, it has several drawbacks that organizational learning approach can improve. First, the goal of managerialism decision making is for better results of standardized test instead of for better learning. As a result, educators' decision makings were inevitably driven by a "teaching to the tests" environment. Second, school administrators' decision makings are mainly for accountability purpose, typically, monitoring the compliance of teachers. Third, localization is important to students learning. Loosely coupled organizations used to protect localization by buffering external changes. Managerialism, however, places external standards as the highest priority for decision making and thereby ignores local needs. Organizational learning may input more flexibility into the system in order to balance the local needs and external standards via constant learning and continuous improvement. Last but not least, decision making under managerialism has a tendency to reinforce traditions, but Organization Learning promotes open communications, and encourages new knowledge to reinvent practice.

Overall, the literature suggested that the managerialism approach and the organizational learning approach have different preferences on data, and different foci in terms of decision making. Therefore, it becomes important to know how "using data" as a coupling mechanism for decision-making works with different paradigm.

Summary

This Chapter reviewed the literature about loose coupling theory, new managerialism and its influence on education reform practice in the U.S., and a few topics related to data-informed decision-making. A theme that runs through the issues in the foregoing literature review is how data-informed decision-making at the various levels is related, particularly when it comes to the

different paradigms of transcending the loosely coupled educational system. In the following chapter, I explain the methodology for the current study.

CHAPTER III

METHODOLOGY

In this chapter, I will address methodological issues including data source and samples, instruments and variables, and statistical procedures.

Data Source and Sample

The data for this study was from a nationally representative 2003-04 School and Staffing Survey (SASS), which was designed by the National Center for Education Statistics (NCES) of the U.S. Department of Education and implemented by the U.S. Census Bureau. NCES provided both public-use SASS data and restricted-use SASS data. The restricted-use data differs from the public-use data in that it contains the identifiable information, which is confidential and protected by federal law. To protect the confidentiality, SASS requires data users from qualified organizations to go through a strict licensing process.

The present study used restricted-use data from the 2003-04 SASS Public Teacher, Principal, School, and School District surveys. The Teacher surveys were designed to collect the teacher information including teacher's demographic information, preparation, teaching experience, professional development, their perceptions and attitudes toward their teaching, and so on. Principal surveys were designed to measure principal's information such as demographic information, preparation, leadership experience, perceptions of their influences and school problems. School surveys have information about school enrollment, grades offered, school programs, school performance and so on. The school district surveys include district information such as total enrollment, student demographic, teacher hiring and evaluation, principal and

teacher salary, and graduation requirements. Overall, SASS data sets offered very comprehensive information of educational practice across all level of education systems.

The 2003-04 SASS Public Teacher, Principal, School, and School District surveys were distributed based on a stratified sampling design, thus each sampled teacher, principal, school, and school district carries a certain weight. The weights were designed to meet various requirements, and one of them is to make sure each sampled teacher, principal, school, and school district would represent those with similar characteristics in respective national populations of teachers, principals, schools, and districts. The present study used a sample of 39,330 public school full-time and regular teachers, 8,140 public school principals/schools, and 4,420 school districts, who after weighting, were representatives of 2,948,230 national public school full-time and regular teachers, 87,620 nationwide public school principals/schools, and 15,540 school districts[1]. Restricted-use SASS allows users to match, link, or even combine different sets of survey data according to specific identification numbers that were assigned to teachers, schools, principals, and districts. In the present study, I merged the SASS data by combining the School data and School District data according to their control number.

Instruments and Variables

Two types of variables were extracted from the 2003-04 SASS Public Teacher Data, which included variables for the composite outcome and five variables for background information. The three variables for the composite outcome measured the extent to which teachers use the information from their students' state or district achievement test scores for their instructional decision-making, which include: (a) to group students into instructional groups by achievement or ability, (b) to assess areas where the teacher needs to strengthen his or her content knowledge or teaching practices, and (c) to adjust the curriculum in areas where his or

her students encountered problems. The measurement scale for these three outcomes ranged from 1= not at all, to 2 = to a small extent, 3 = to a moderate extent, and 4 = to a great extent. These three variables indicate teacher's use of student test scores for different instructional decisions. I am interested in an index—teacher's overall use of student test scores for all the three instructional decisions. An index, according to Babbie (2010), is constructed through "accumulating scores assigned to individual attributes" (p. 162). The three variables had high internal consistency (Cronbach's alpha = .991). Thus, a composite index of teachers' data-informed instructional decision-making (TDIDM)—was created as the average of the three outcomes. The five teacher background variables, for control purposes, included gender, certificate level, teaching level, teaching assignment, and teaching experience. The school average of teachers' three instructional decision areas, gender, certificate level, and teaching experience were further calculated as school level predictors in study 3.

From the 2003-04 SASS Public Principal Data, three key variables were extracted to measure whether the schools use any of the following to assess their progress with the formal school improvement plan: (a) state or national tests (Test) (0 = no, 1 = yes), (b) parent or student surveys (Survey) (0 = no, 1 = yes), and (c) student portfolio (Portfolio) (0 = no, 1 = yes). These variables were utilized in two approaches. First, based on these three variables, a composite index was constructed. I checked the internal consistency (Cronbach's alpha = .615), and constructed a composite index variable (PDS) to measure the sources of data available to the principals (0 = none, 1= used one of the three, 2 = used two of the three, and 3 = used all three). Second, they were used discretely for identifying their specific effects on the extent of teachers' data-informed decision-making. Also from this data set, for control purposes, seven principals' and schools' background variables were extracted, including gender, experience, school level,

school enrollment, percentage of minority students, percentage of students enrolled in national free and reduced lunch program, percent of teachers teaching to high academic standards, and school performance in terms of passing district and state performance. For study 1, school performance was used as a control variable; For study 3, school performance was used as an outcome variable (see Appendix A for more information about coding scheme).

Another five key variables were extracted from 2003-04 SASS Public School Data, which measured the extent of schools' application of student performance reports for decision making. The decision making types included: (a) evaluating the progress of students in the school, (b) determining the next year's instructional focus, (c) realigning the curriculum, such as with content standards and/or other indicator criteria, (d) informing parents and the community of the school's progress, and (e) prompting school-level initiatives for improvement. The measure for the five variables were all binary (0 = no, 1 = yes). Similarly, these five variables were used through two approaches. First, based on these five variables, a composite index was constructed. I checked both the tetrachoric correlations among the five variables, which ranged from .781 to .884, and the internal consistency (Cronbach's alpha = .825), and constructed a composite index to measure the extent of schools' data-informed decision-making (SDIDM) with a focus on types of decisions made (0 = none; 1 = 1 type; 2 = 2 types; 3 = 3 types; 4 = 4 types; and 5 = 5 types). Second, they were used discretely for identifying their specific effects on the extent of teachers' data-informed decision-making. These variables were merged with principal data and used as the second-level variables for the statistical analysis.

From the 2003-04 SASS School District data, I extracted five variables which were essentially the same as the five key variables from the School Data. The decision making types included (a) evaluating the progress of students in the school, (b) determining the next year's

instructional focus, (c) realigning the curriculum, such as with content standards and/or other indicator criteria, (d) informing parents and the community of the school's progress, and (e) prompting school-level initiatives for improvement. The measure for the five variables were also binary (0 = no, 1 = yes). These five items were used for identifying the effects of district's types of using data on the extent of school's data-informed decision-making (Cronbach's alpha = .913).

Finally, I extracted weighted variables for all three data sets and the linking variable for the teacher, principal, and district data sets (i.e., the school control number; the district control number). I applied relative weights at all three levels to make the teacher, principal, and district samples nationally representative while keeping the original sample sizes so as not to inflate the statistics. The HLM software allows the application of weights at all levels simultaneously. Please see appendix A for the name, measurement scale, and descriptive statistics for the variables used in the study.

Statistical Procedures

HLM designs were proven to be effective and reliable in organizational analysis and policy evaluation (see Chatterji, 2002; Puma et al., 2000). Considering the hierarchical structure of educational data, and the inadequacy of traditional statistical techniques for modeling hierarchical data (Raudenbush & Bryk, 2002), two-level HLM were applied in the 3 studies.

Statistical Procedure for Research Study 1

(a) To what extent is teachers' data-informed decision-making attributable to the school level? (b) Is there a significant variation among schools in teachers' data-informed decision-making? And (c) after controlling for teachers', principals', and schools' background variables,

are principals' and schools' data-informed decision-making associated with teachers' data-informed decision-making? If so, is it a pattern that inspired by accountability purpose or organizational learning culture?

Specifically, teachers were at the first level, and schools/principals were at the second level for Question 1. The independent variables at first level were estimated as fixed effects, meaning that all slopes were constant across schools. In spite of the high correlation between some variables, no variance inflation factors (VIFs) were greater than 4 for variables at the first and second levels, respectively. Therefore, multicollinearity was not a concern. Further, a four-step model building strategy was implemented to facilitate the analysis. I used the robust standard errors for the final estimation of fixed effects, an approach that is not sensitive to violation of assumptions of homogeneity and normality given our large sample size (Cheong, Fotiu, & Raudenbush, 2001; Hox & Maas, 2001; Kauermann & Carroll, 2001, Wang & Long, 2011).

Step 1: Unconditional Model. The first step was to build Model 1, an unconditional model without any predictor, which functioned to gauge to what extent a second level model could contribute to explaining the variance of the outcome variable. The first level model of Model 1 is:

$$(TDIDM)_{ij} = \beta_{0j} + r_{ij}$$
$$r_{ij} \sim N(0, \sigma^2) \qquad (1.1)$$

Where TDIDM was teacher i's data-informed decision-making in school j, β_{0j} was the teachers' average data-informed decision-making in school j, r_{ij} was the teacher level random effect, and σ^2 was the variance within schools.

The second level model of Model 1 is:

$$\beta_{0j} = \gamma_{00} + \mu_{0j} \qquad (1.2)$$
$$\mu_{0j} \sim N(0, \tau_{00})$$

Where γ_{00} was the grand mean of all teachers' data-informed decision-making across all national public schools, μ_{0j} was the principal/school level random effect, and τ_{00} was the variance across all national public schools.

To gauge to what extent a second level model could contribute to explaining the variance of the outcome variable, we calculated the intra-class correlation by following the equation:

$$\rho = \frac{\hat{\tau}_{00}}{\hat{\sigma}^2 + \hat{\tau}_{00}}$$

Where $\hat{\sigma}^2$ and $\hat{\tau}_{00}$ were the estimates of level-1 and level-2 variances.

Step 2: Control Models for Model Comparison. In order to capture the variance explained by variables at different levels, Raudenbush and Bryk (2002) suggested building the first level first. Therefore, before building the full models, two control models needed to be established. The first control model—Model 2 added only teachers' background variables at the first level, while the second level remained the same as in Model 1:

$$(TDIDM)_{ij} = \beta_{0j} + \beta_{1j}(Gender) + \beta_{2j}(Cert) + \beta_{3j}(Level)$$
$$+\beta_{4j}(Field) + b_{5j}(Experience) + r_{ij} \qquad (2)$$
$$r_{ij} \sim N(0, \sigma^2)$$

Where $\beta_{1j} - \beta_{5j}$ were the effects of the five teacher background variables and r_{ij} was the teacher level random effect with variance σ^2.

The second control model, Model 3 added the principals'/schools' background variables to Model 2 at the second level while level 1 remained the same as in Model 2:

$$\begin{aligned}\beta_{0j} &= \gamma_{00} + \gamma_{01}(Gender) + \gamma_{02}(Experience) + \gamma_{03}(SchLevel1) \\ &+\gamma_{04}(SchLevel2) + \gamma_{05}(Enrollment) + \gamma_{06}(Minority) + \gamma_{07}(NSLP) \\ &+\gamma_{08}(Location1) + \gamma_{09}(Location2) + \gamma_{010}(SchPerformance) + \mu_{0j} \\ \mu_{0j} &\sim N(0, \tau_{00})\end{aligned} \quad (3)$$

Where γ_{01}– γ_{010} were the effects of principals'/schools' background variables.

These two control models could help identify the model improvement. For instance, the proportion of variance explained at the first level of Model 2 over Model 1 was

$$Proportion\ of\ \ variance\ explained = \frac{\sigma_{00}(\text{Model 1}) - \sigma_{00}(\text{Model 2})}{\sigma_{00}(\text{Model 1})}$$

Similarly, the proportion of variance explained at the second level of Model 3 over Model 2 was

$$Proportion\ of\ \ variance\ explained = \frac{\sigma_{00}(\text{Model 2}) - \sigma_{00}(\text{Model 3})}{\sigma_{00}(\text{Model 2})}$$

Where $\sigma_{00}(\text{Model 1})$, $\sigma_{00}(\text{Model 2})$, and $\sigma_{00}(\text{Model 3})$ were the estimates of level-2 variances of Model 1, Model 2, and Model 3.

Step 3: Full Models for the Composite Variables. Based on the control models, another two steps were applied to investigate the effects of principals' and schools' utilization of various data and using data for decision making. First, two models were utilized to model the effects of the two composite variables: PDS (i.e., type of data used) and SDIDM (i.e., types of decisions made).

Model 4 added principals' use of different types of data, PDS, to the second-level of Model 3:

$$\begin{aligned}\beta_{0j} &= \gamma_{00} + \gamma_{01}(Gender) + \gamma_{02}(Experience) + \gamma_{03}(SchLevel1) + \gamma_{04}(SchLevel2) \\ &+\gamma_{05}(Enrollment) + \gamma_{06}(Minority) + \gamma_{07}(NSLP) + \gamma_{08}(Location1) \\ &+\gamma_{09}(Location2) + \gamma_{010}(SchPerformance) + \gamma_{011}(PDS) + \mu_{0j} \\ \mu_{0j} &\sim N(0, \tau_{00})\end{aligned} \quad (4)$$

Where γ_{011} represented the effect of the extent of principals' utilization of data (i.e., type of data used) on teachers' data-informed decision-making.

Model 5 added SDIDM to the second level of Model 4.

$$\begin{aligned}\beta_{0j} &= \gamma_{00} + \gamma_{01}(Gender) + \gamma_{02}(Experience) + \gamma_{03}(SchLevel1) + \gamma_{04}(SchLevel2) \\ &+\gamma_{05}(Enrollment) + \gamma_{06}(Minority) + \gamma_{07}(NSLP) + \gamma_{08}(Location1) \\ &+\gamma_{09}(Location2) + \gamma_{010}(SchPerformance) + \gamma_{011}(PDS) + \gamma_{012}(SDIDM) + \mu_{0j} \\ \mu_{0j} &\sim N(0, \tau_{00})\end{aligned} \quad (5)$$

Where γ_{012} represented the effect of the extent of schools' application of student performance report for decision making on teachers' data-informed decision-making.

In order to know how much variance had been explained by adding PDS to Model 3 and by adding SDIDM to Model 4, we calculated the proportion of additional variance explained. According to Raudenbush and Bryk (2002), when the level-1 model did not change, then the level-1 variance remained the same, while level-2 variance would decrease when significant predictors were included. Thus, after adding PDS to Model 3, the proportion of additional variance explained by PDS could be computed by the following equation:

$$Proportion\ of\ \ variance\ explained = \frac{\sigma_{00}(\text{Model 3}) - \sigma_{00}(\text{Model 4})}{\sigma_{00}(\text{Model 3})}$$

Similarly, after adding SDIDM to Model 4, the proportion of extra variance explained was:

$$Proportion\ of\ \ variance\ explained = \frac{\sigma_{00}(\text{Model 4}) - \sigma_{00}(\text{Model 5})}{\sigma_{00}(\text{Model 4})}$$

Where $\sigma_{00}(\text{Model 3})$, $\sigma_{00}(\text{Model 4})$, and $\sigma_{00}(\text{Model 5})$ were the estimates of level-2 variances of Model 3, Model 4, and Model 5.

Step 4: Full Models for the Specific Variables. To investigate the specific effect of types of school data used and schools' application of student performance report for decision making,

the second step was to replace the two composite variables with the detailed information about principals' and schools' data-informed decision-making (three types of school data used, and five types of data-informed decision-making).

Model 6 added the three types of school data used (Test, Survey, and Portfolio) to Model 3 at level-2:

$$\begin{aligned}
&\beta_{0j} = \gamma_{00} + \gamma_{01}(Gender) + \gamma_{02}(Experience) + \gamma_{03}(SchLevel1) + \gamma_{04}(SchLevel2) \\
&+\gamma_{05}(Enrollment) + \gamma_{06}(Minority) + \gamma_{07}(NSLP) + \gamma_{08}(Location1) + \gamma_{09}(Location2) \\
&+\gamma_{010}(SchPerformance) + \gamma_{011}(Test) + \gamma_{012}(Survey) + \gamma_{013}(Portfolio) + \mu_{0j} \\
&\mu_{0j} \sim N(0, \tau_{00})
\end{aligned} \tag{6}$$

Where γ_{011} - γ_{013} represented the effect of principals' utilization of the three types of data (Test, Survey, and Portfolio) on teachers' data-informed decision-making.

Model 7 added the five types of schools' data-informed decision-making to Model 6 at level-2:

$$\begin{aligned}
&\beta_{0j} = \gamma_{00} + \gamma_{01}(Gender) + \gamma_{02}(Experience) + \gamma_{03}(SchLevel1) + \gamma_{04}(SchLevel2) \\
&+\gamma_{05}(Enrollment) + \gamma_{06}(Minority) + \gamma_{07}(NSLP) + \gamma_{08}(Location1) + \gamma_{09}(Location2) \\
&+\gamma_{010}(SchPerformance) + \gamma_{011}(Test) + \gamma_{012}(Survey) + \gamma_{013}(Portfolio) \\
&+\gamma_{014}(Evaluating\ students) + \gamma_{015}(Determining\ focus) + \gamma_{016}(Realigning\ curriculum) \\
&+\gamma_{017}(Informing\ parents) + \gamma_{018}(Prompting\ initiatives) + \mu_{0j} \\
&\mu_{0j} \sim N(0, \tau_{00})
\end{aligned} \tag{7}$$

Where γ_{014} - γ_{018} represented the effect of schools' application of data-informed decision-making in the five areas on teachers' data-informed decision-making.

Likely, after adding the three specific principals' utilization of school data to Model 3, the proportion of additional variance explained could be computed by the following equation:

$$Proportion\ of\ \ variance\ explained = \frac{\sigma_{00}(\text{Model 3}) - \sigma_{00}(\text{Model 6})}{\sigma_{00}(\text{Model 3})}$$

Similarly, after adding the five types of schools' data-informed decision-making to Model 6, the proportion of additional variance explained was:

$$Proportion\ of\ \ variance\ explained = \frac{\sigma_{00}(\text{Model 6}) - \sigma_{00}(\text{Model 7})}{\sigma_{00}(\text{Model 6})}$$

Where $\sigma_{00}(\text{Model 3})$, $\sigma_{00}(\text{Model 6})$, and $\sigma_{00}(\text{Model 7})$ were the estimates of level-2 variances of Model 3, Model 6, and Model 7.

Statistical Procedure for Research Study 2

For the second study, I have the following questions: (a) To what extent is schools' data-informed decision-making attributable to the district level? (b) Is there a significant variation among districts in schools' data-informed decision-making? And (c) after controlling for principals', schools', and districts' background variables are districts' data-informed decision-making associated with schools' data-informed decision-making? If so, is it a pattern that inspired by accountability purpose or organizational learning culture?

Given the similar data structure and similar sets of research questions, statistical procedures from study 1 were applied to the study 2. However, the school level was placed at the first level, and district level was placed at the second level.

Step 1: Unconditional Model. The first step was to build Model 1, an unconditional model without any predictor to gauge to the variance of the outcome variable explained by second level. The first level model of Model 1 is:

$$\begin{aligned} (SDIDM)_{ij} &= \beta_{0j} + r_{ij} \\ r_{ij} &\sim N(0, \sigma^2) \end{aligned} \qquad (1.1)$$

Where SDIDM was school i's data-informed decision-making in district j, β_{0j} was the schools' average data-informed decision-making in district j, r_{ij} was the school level random effect, and σ^2 was the variance within districts.

The second level model of Model 1 is:

$$\beta_{0j} = \gamma_{00} + \mu_{0j} \qquad (1.2)$$
$$\mu_{0j} \sim N(0, \tau_{00})$$

Where γ_{00} was the grand mean of all schools' data-informed decision-making across all national public districts, μ_{0j} was the district level random effect, and τ_{00} was the variance across all national public districts.

Step 2: Control Models for Model Comparison. The first control model—Model 2 added only school and principal background variables at the first level, while the second level remained the same as in Model 1:

$$\begin{aligned}(SDIDM)_{ij} &= \beta_{0j} + \beta_{1j}(Gender) + \beta_{2j}(Experience) + \beta_{3j}(SchLevel1) \\ &+ \beta_{4j}(SchLevel2) + \beta_{5j}(Enrollment) + \beta_{6j}(Minority) + \beta_{7j}(NSLP) \\ &+ \beta_{8j}(Location1) + \beta_{9j}(Location2) + \beta_{10j}(SchPerformance) + \beta_{11j}(PDS) + r_{11j} \\ r_{ij} &\sim N(0, \sigma^2)\end{aligned} \qquad (2)$$

Where $\beta_{1j} - \beta_{10j}$ were the effects of the school background variables and r_{ij} was the school level random effect with variance σ^2.

Model 3 added the principals'/schools' background variables to Model 2 at the second level while level 1 remained the same as in Model 2:

$$\begin{aligned}\beta_{0j} &= \gamma_{00} + \gamma_{01}(DNSLP) + \gamma_{02}(DEnrollment) + \mu_{0j} \\ \mu_{0j} &\sim N(0, \tau_{00})\end{aligned} \qquad (3)$$

Where $\gamma_{01} - \gamma_{010}$ were the effects of principals'/schools' background variables.

The proportion of variance explained at the first level of Model 2 over Model 1 was

$$Proportion\ of\ \ variance\ explained = \frac{\sigma_{00}(\text{Model 1}) - \sigma_{00}(\text{Model 2})}{\sigma_{00}(\text{Model 1})}$$

The proportion of variance explained at the second level of Model 3 over Model 2 was

$$Proportion\ of\ \ variance\ explained = \frac{\sigma_{00}(\text{Model 2}) - \sigma_{00}(\text{Model 3})}{\sigma_{00}(\text{Model 2})}$$

Where $\sigma_{00}(\text{Model 1})$, $\sigma_{00}(\text{Model 2})$, and $\sigma_{00}(\text{Model 3})$ were the estimates of level-2 variances of Model 1, Model 2, and Model 3.

Step 3: Full Models for the Composite Variables. Based on the control models, another two steps were applied to investigate the effects of districts' utilization of various data sources and using data for decision making. First, two models were utilized to model the effects of the DDIDM (i.e., types of decisions made at district level).

Model 4 added data types available in the districts, DDIDM, to the second-level of Model 3:

$$\begin{aligned} &\beta_{0j} = \gamma_{00} + \gamma_{01}(DNSLP) + \gamma_{02}(DEnrollment) + \gamma_{03}(DDIDM) + \mu_{0j} \\ &\mu_{0j} \sim N(0, \tau_{00}) \end{aligned} \qquad (4)$$

Where γ_{03} represented the effect of the extent of district data-informed decision (i.e., type of data used) on schools' data-informed decision-making.

After adding DDIDM to Model 3, the proportion of additional variance explained by DDIDM could be computed by the following equation:

$$Proportion\ of\ \ variance\ explained = \frac{\sigma_{00}(\text{Model 3}) - \sigma_{00}(\text{Model 4})}{\sigma_{00}(\text{Model 3})}$$

Where $\sigma_{00}(\text{Model 3})$ and $\sigma_{00}(\text{Model 4})$ were the estimates of level-2 variances of Model 3 and Model 4.

Step 4: Full Models for the Specific Variables. To investigate the specific effect of types of district data used and districts' application of student performance report for decision making, the two composite variables were replaced with the detailed items (ten types of district data sources, and five types of data application).

Model 5 added the five types of schools' data-informed decision-making to Model 3 at level-2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(DNSLP) + \gamma_{02}(DEnrollment) + \gamma_{03}(Devaluating\ students) + \gamma_{04}(Ddetermining\ focus) + \gamma_{05}(Drealigning\ curriculum) + \gamma_{06}(Dinforming\ parents) + \gamma_{07}(Dprompting\ initiatives) + \mu_{0j}$$
$$\mu_{0j} \sim N(0, \tau_{00}) \quad (5)$$

Where γ_{03} - γ_{07} represented the effect of districts' application of data-informed decision-making in the five areas on school's data applications.

The proportion of additional variance explained could be computed by the following equation:

$$Proportion\ of\ variance\ explained = \frac{\sigma_{00}(\text{Model 3}) - \sigma_{00}(\text{Model 5})}{\sigma_{00}(\text{Model 3})}$$

Where $\sigma_{00}(\text{Model 3})$ and $\sigma_{00}(\text{Model 5})$ were the estimates of level-2 variances of Model 3 and Model 5.

Statistical Procedure for Research Study 3

In the study 3, the question is to what extent is school performance attributable to data-informed decision-making at two levels?

Step 1: Unconditional Model. Similar to Question 1 and Question 2, the first step was to build an unconditional model without any predictor. However, unlike study 1 and study 2, which had continuous outcomes, study 3 had a dichotomous outcome variable – if principals reported that their schools did not pass/pass some/pass most of the district or state standards, the school performance was coded as 0; if principals reported that their schools passed all district or state standards, their performance was then coded as 1. Therefore, a non-linear HLM was employed, where the outcome value was transformed with a logit link function. The level-1 model was:

$$Y_{ij} \mid \varphi_{ij} \sim B(m_{ij}, \varphi_{ij})$$

The level-1 link function was:

$$\eta_{ij} = \log\left(\frac{\phi_{ij}}{1-\phi_{ij}}\right) = \log\left(\frac{\text{Prob}(SchPass = 1)}{1-\text{Prob}(SchPass = 1)}\right)$$

Where ϕ_{ij} was the probability of school passing all standards, and η_{ij} was the log odds of whether schools passed all standards within district j.

The level-1 model of Model 1 is:

$$\eta_{ij} = \beta_{0j} \qquad (1.1)$$

And level-2 model is:

$$\beta_{0j} = \gamma_{00} + u_{0j} \qquad (1.2)$$

$$u_{0j} \sim N(0, \tau_{00})$$

Where γ_{00} is the average log-odds of passing all standards across all districts. And τ_{00} is the variance between districts in district-average log-odds of passing all standards. The intra-class correlation of models with binary outcomes could be calculated with following the equation:

$$\rho = \frac{\tau}{\pi^2 / 3 + \tau}$$

Where τ was the estimate of level-2 variances.

Step 2: Control Models for Model Comparison. Two control models were established. The first control model—Model 2 added only school's background variables at the first level, while the second level remained the same as in Model 1:

$$\eta_{ij} = \beta_{0j} + \beta_{1j}(Gender) + \beta_{2j}(Experience) + \beta_{3j}(SchLevel1) + \beta_{4j}(SchLevel2)$$
$$+\beta_{5j}(Enrollment) + \beta_{6j}(Minority) + \beta_{7j}(NSLP) + \beta_{8j}(Location1) + \beta_{9j}(Location2) \quad (2)$$
$$+\beta_{10j}(PDS) + \beta_{11j}(TSexp) + \beta_{12j}(HighTeach) + \beta_{13j}(TScert) + \beta_{14j}(TSgender)$$

Where $\beta_{1j} - \beta_{14j}$ were the effects of the principals'/schools' background variables.

The second control model, Model 3 added the district background variables to Model 2 at the second level while level 1 remained the same as in Model 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(DNSLP) + \gamma_{02}(DEnrollment) + \mu_{0j} \quad (3)$$
$$\mu_{0j} \sim N(0, \tau_{00})$$

Where $\gamma_{01} - \gamma_{02}$ were the effects of districts' background variables. These two control models could help identify the model improvement.

Step 3: Full Models for the Composite Variables without Cross-level effects. Based on the control models, another step was applied to investigate the effects of using data for decision making at school level and district level on school performance. Three specific school level teacher average data-informed decision-making – using test data to group student, to assess areas to strengthen, and to adjust curriculum were added to level 1 in Model 3.

Model 4:

$$\eta_{ij} = \beta_{0j} + \beta_{1j}(Gender) + \beta_{2j}(Experience) + \beta_{3j}(SchLevel1) + \beta_{4j}(SchLevel2)$$
$$+\beta_{5j}(Enrollment) + \beta_{6j}(Minority) + \beta_{7j}(NSLP) + \beta_{8j}(Location1) + \beta_{9j}(Location2)$$
$$+\beta_{10j}(PDS) + \beta_{11j}(TSexp) + \beta_{12j}(HighTeach) + \beta_{13j}(TScert) + \beta_{14j}(TSgender) \quad (4)$$
$$+\beta_{15j}(GroupStudent) + \beta_{16j}(AssessAreas) + \beta_{17j}(AdjustCurriculum)$$

Where $\beta_{15j} - \beta_{17j}$ represent the effect of three specific school level teacher average data-informed decision-making on school performance.

Five specific school decision types were added to level 1 in Model 4, while the level 2 still remained the same as Model 2.

Model 5:

$$\eta_{ij} = \beta_{0j} + \beta_{1j}(Gender) + \beta_{2j}(Experience) + \beta_{3j}(SchLevel1) + \beta_{4j}(SchLevel2)$$
$$+\beta_{5j}(Enrollment) + \beta_{6j}(Minority) + \beta_{7j}(NSLP) + \beta_{8j}(Location1) + \beta_{9j}(Location2)$$
$$+\beta_{10j}(PDS) + \beta_{11j}(TSexp) + \beta_{12j}(HighTeach) + \beta_{13j}(TScert) + \beta_{14j}(TSgender)$$
$$+\beta_{15j}(GroupStudent) + \beta_{16j}(AssessAreas) + \beta_{17j}(AdjustCurriculum) \quad (5)$$
$$+\beta_{18j}(Evaluating\ students) + \beta_{19j}(Determining\ focus) + \beta_{20j}(Realigning\ curriculum)$$
$$+\beta_{21j}(Informing\ parents) + \beta_{22j}(Prompting\ initiatives)$$

Lastly, five district level data-informed decision types were added to level 2 in Model 4.

Model 6:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(DNSLP) + \gamma_{02}(DEnrollment) + \gamma_{03}(Devaluating\ students)$$
$$+\gamma_{04}(Ddetermining\ focus) + \gamma_{05}(Drealigning\ curriculum)$$
$$+\gamma_{06}(Dinforming\ parents) + \gamma_{07}(Dprompting\ initiatives) + \mu_{0j} \quad (6)$$
$$\mu_{0j} \sim N(0, \tau_{00})$$

Where $\gamma_{03} - \gamma_{06}$ represented the effect of the extent of average districts' utilization of data on school performance. Again, the proportion of additional variance explained by these composite data-informed decision-making variables was calculated.

Summary

This chapter explained the data sources, instruments, and statistical procedures of the present study. A group of teacher items, school items, principal items, and district items were selected and recoded from SASS 2003-2004 surveys. Data extracted from those items were matched according to control ID of different level and analyzed with HLM statistical procedures to answer three sets of research questions that were raised in Chapter 1. The following chapter, Chapter 4 will present the statistical findings based on the methodology reported in this chapter.

CHAPTER IV

FINDINGS

Findings of Study 1

Variance of TDIDM

After conducting the analysis for the unconditional model, the results were presented in Table 2. The results showed that the level-1 variance was 1.32584, while the level-2 variance was 0.22253. Thus, the intra-class correlation (ICC) was 0.22253/ (1.32584 + 0.22253) = 0.144. Therefore, the finding indicated that adding a level-2 model can potentially explain 14.4% of the outcome variance. In other words, 14.4% of the variability in teachers' data-informed decision-making exists between schools. We expected to see some variance to be explained at both levels when we added teacher characteristics and principal/school characteristics to Model 1.

The results of first control model (Model 2) were included in Table 2 as well. As we can see, all five teacher characteristics were statistically significant. The conditional ICC is 0.19315 / (1.30985 + 0.19315) = 0.129, indicating that after controlling for gender, certificate, teaching level, teaching assignment, and teaching experience, we still had 12.9% of the outcome variability existing between schools. And the level-2 variance was still significant ($p < .001$) showing that the variance was expected to be explained by adding potential predictors.

Table 2

Results of the Unconditional Model and the First Control Model of TDIDM (with robust standard errors)

Fixed Effect	Model 1 Coefficient	S.E.	Model 2 Coefficient	S.E.
Intercept	2.233***	0.013	2.368***	0.056
teacher gender			-0.239***	0.023
teacher certificate			0.151***	0.040
teaching level			-0.239***	0.025
teaching assignment			0.202***	0.024
teaching experience			0.008***	0.001
Random Effect	Variance	*p*	Variance	*p*
Level-2 between school	0.22253	0.000	0.19315	0.000
Level-1 within school	1.32584		1.30985	
Intra-class correlation	0.144		0.129	

Note: ***p < .001;

Effects of Composite PDS and Composite SDIDM on TDIDM

The results of Model 3, Model 4, and Model 5 were included in Table 3. The results of Model 3 revealed seven significant principal/school characteristics: principal gender, school level-elementary, school level-secondary, school total enrollment, percentage of minority students, percentage of students enrolled in the national free and reduced lunch program, and school location. Compared with Model 2, the improvement of Model 3 at level-2 could be gauged by the proportion of variance explained: (0.19315 – 0.15347) / 0.19315 = 0.205. This indicated that after adding the eight principal/school characteristics, level-2 variance was reduced by 20.5%, which was a large amount.

The results of Model 4 showed that the seven principal/school characteristics were still significant. And most importantly, the focused composite variable—PDS was significant and

positive (β= 0.056, p < .001, effect size[1]= 0.05). The improvement of Model 4 at level-2 could also be gauged by the proportion of variance explained: (0.15347 – 0.15046) / 0.15347 = 0.02. This indicated that after including PDS, level-2 variance was reduced by 2%. Overall the results suggested that there was a statistically significant and positive relationship between principals' data-informed decision-making and teachers' data-informed decision-making, but the relationship tended to be weak as measured by the proportion of variance explained and the effect size for the coefficient.

The results of Model 5 showed that the composite variable SDIDM was significant and positive (β= 0.031, p < .001, effect size = 0.03), while PDS was still significant (β= 0.054, p < .001, effect size = 0.04). All other school/principal characteristics remain either significant or insignificant as in Model 4. The improvement of Model 5 at level-2 was gauged by the proportion of variance explained: (0.15046 – 0.14964) / 0.15046 = 0.005. This indicated that after including PDS, level-2 variance was reduced by .5%. Overall we found that both principals' and schools' data-informed decision-making (as measured by PDS and SDIDM) had statistically significant and positive effects on teachers' data-informed decision-making, but the relationships tended to be weak as measured by the proportion of variance explained and the effect sizes for the coefficients.

[1] The effect size for a coefficient in hierarchical linear modeling can be calculated with equation $\frac{\beta}{\sqrt{\tau^2 + \sigma^2}}$, where β is the coefficient, τ^2 is the between-school variance from the unconditional model, and σ^2 is within-school variance from the unconditional model (see McCoach, 2010). Rosenthal and Rosnow (1984) classified effect sizes more than .50 as large, between .30 and .50 as moderate, and less than .30 as small.

Table 3

Results of the Impact of PDS and SDIDM on TDIDM (with robust standard errors)

Fixed Effect	Model 3		Model 4		Model 5	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Intercept	1.810***	0.086	1.713***	0.089	1.584***	0.109
level-1						
teacher gender	-0.224***	0.023	-0.222***	0.023	-0.223***	0.023
teacher certificate	0.165***	0.040	0.166***	0.040	0.166***	0.041
teaching level	-0.002	0.033	0.001	0.033	0.000	0.033
teaching assignment	0.191***	0.024	0.191***	0.024	0.191***	0.024
teaching experience	0.009***	0.001	0.009***	0.001	0.009***	0.001
level-2						
principal gender	-0.061*	0.028	-0.056*	0.028	-0.055	0.028
principal experience	-0.001	0.002	-0.002	0.002	-0.002	0.002
school level-elementary	0.202***	0.047	0.188***	0.046	0.173***	0.045
school level-secondary	-0.094*	0.045	-0.105*	0.045	-0.113*	0.044
school enrollment	-0.008**	0.002	-0.009**	0.002	-0.009***	0.002
percent of minority student	0.002**	0.001	0.001*	0.001	0.001*	0.001
percent of NSLP student	0.004***	0.001	0.004***	0.001	0.004***	0.001
school location-urban	-0.151***	0.041	-0.147**	0.041	-0.143**	0.041
school location-suburban	-0.058	0.031	-0.053	0.030	-0.051	0.030
school performance	-0.032	0.025	-0.034	0.025	-0.035	0.025
PDS			0.056***	0.013	0.054***	0.013
SDIDM					0.031*	0.014
Random effects	Variance	*p*	Variance	*p*	Variance	*p*
Level-2	0.15347	0.000	0.15046	0.000	0.14964	0.000
Level-1	1.31473		1.31518		1.31537	
Conditional ICC	0.105		0.103		0.102	
	Over Model 2		Over Model 3		Over Model 4	
Proportion of variance explained at level-2	0.205		0.020		0.005	

Note: *** $p < .001$; ** $.001 < p < .01$; * $.01 < p < .05$

Effects of Specific PDS and Specific SDIDM on TDIDM

The results of Model 6 and Model 7 were included in Table 4. For the purpose of model comparison, Model 3 was also included as in Table 4. After adding principals' utilization of three data type (Test, Portfolio, and Survey) to Model 3, the results of Model 6 were presented in the fourth and fifth columns. The results showed that among the utilization of data types, the only significant type was Test ($\beta = 0.107$, $p < .05$, effect size = 0.09), while Portfolio and Survey were not significant ($p > .05$). The findings suggested that principals' use of test scores had a positive effect on teachers' classroom-level practice, while principals' use of data from portfolio and surveys did not. Comparing with Model 3, the improvement of Model 6 at level-2 was gauged by the proportion of variance explained: (0.15347 – 0.15029) / 0.15347 = 0.021. This suggested that after including the three individual school utilizations of data type, level-2 variance was reduced by 2.1%.

The results of Model 7 showed that, among these five new predictors, the only significant variable was schools' application of student performance report for determining the next year's instructional focus (β= 0.122, $p < .05$, effect size = 0.10), while Test was still significant (β= 0.098, $p < .001$, effect size = 0.08). The findings suggested that schools' application of student performance report for determining the next year's instructional focus had a positive effect on teachers' classroom-level practice, while schools' use of data from portfolios and surveys did not. The improvement of Model 7 at level-2 was gauged by the proportion of variance explained: (0.15029 – 0.14917) / 0.15029 = 0.007. This indicated that after including the five applications of student performance for decision making, level-2 variance was reduced by .7%.

Table 4

Results of the Impact of Specific PDS and Specific SDIDM on TDIDM (with robust standard errors)

Fixed Effect	Model 3		Model 6		Model 7	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Intercept	1.810***	0.086	1.689***	0.090	1.577***	0.114
level-1						
teacher gender	-0.224***	0.023	-0.222***	0.023	-0.223***	0.023
teacher certificate	0.165***	0.040	0.166***	0.040	0.165***	0.040
teaching level	-0.002	0.033	0.000	0.033	0.000	0.033
teaching assignment	0.191***	0.024	0.191***	0.024	0.191***	0.024
teaching experience	0.009***	0.001	0.009***	0.001	0.009***	0.001
level-2						
principal gender	-0.061*	0.028	-0.058*	0.028	-0.056*	0.028
principal experience	-0.001	0.002	-0.002	0.002	-0.002	0.002
school level-elementary	0.202***	0.047	0.184***	0.046	0.167***	0.044
school level-secondary	-0.094*	0.045	-0.108*	0.045	-0.114**	0.043
school enrollment	-0.008**	0.002	-0.009**	0.002	-0.009***	0.002
percent of minority student	0.002**	0.001	0.001*	0.001	0.001*	0.001
percent of NSLP student	0.004***	0.001	0.004***	0.001	0.004***	0.001
school location-urban	-0.151***	0.041	-0.146**	0.041	-0.142**	0.041
school location-suburban	-0.058	0.031	-0.052	0.030	-0.050	0.030
school performance	-0.032	0.025	-0.036	0.026	-0.036	0.026
School-level Data-type						
Test			0.107*	0.046	0.098*	0.046
Portfolio			0.026	0.027	0.024	0.027
Survey			0.053	0.034	0.052	0.034

Table 4—Continued

School-level data-informed decision areas						
Evaluating students					0.005	0.068
Determining focus					0.122*	0.053
Realigning curriculum					-0.032	0.050
Informing parents					0.005	0.093
Prompting initiatives					0.046	0.067
Random effects	Variance	*p*	Variance	*p*	Variance	*p*
Level-2	0.15347	0.000	0.15029	*0.000*	0.14917	*0.000*
Level-1	1.31473		1.31518		1.31551	
	Over Model 2		Over Model 3		Over Model 6	
Proportion of variance explained at level-2	.205		.021		.007	

Note: *** p < .001; ** .001 < p < .01; * .01 < p < .05;

In addition, although not the focus of this study, some teacher- and school-level background variables were found to be statistically significant, including those associated with teacher-level instructional data-informed decision-making. At the teacher level, female teachers, higher teaching certificate level, math and science teachers, and more years of teaching were associated with higher level of teachers' instructional data-informed decision-making at $p < .001$. At the school level, elementary school (vs middle and high school) and percentage of students on free and reduced-price lunch were positively—while school enrollment was negatively—associated with teacher-level instructional data-informed decision-making, all at $p < .001$. These findings suggest that certain teacher and school characteristics are related to teacher-level data-informed decision-making. Given the increasing emphasis on data-informed decision-making as a way to improve education, these findings provide some guidance, based on a large nationally representative data set, on facilitators and barriers related to classroom-level data-informed decision-making. Those variables with statistically significant positive or negative correlations with teacher-level data-informed decision-making are facilitators and barriers, respectively.

Findings of Study 2

Variance of SDIDM

After conducting the analysis for the unconditional model, the results were presented in Table 5. The results showed that the level-1 variance was 0.01118, and the level-2 variance was 0.04152. Therefore, the intra-class correlation was 0.04152/ (0.04152+ 0.01118) = 0.7878, suggesting that adding a level-2 model can potentially explain 78.8% of the outcome variance. In

other words, 78.8% of the variability in school's data-informed decision-making exists between districts.

The results of first control model (Model 2) were also included in Table 5. As we can see, seven school/principal characteristics were statistically significant. The conditional ICC is 0.04050/ (0. 04050+ 0. 01095) = 0.7871, indicating that after controlling for principal gender, experience, school level, school enrollment, school location, principal data sources and school performance, 78.7% of the outcome variability existing between districts. And the level-2 variance was still significant ($p < .001$) indicating that the variance was expected to be explained by adding potential predictors.

Table 5

Results of the Unconditional Model and the First Control Model of SDIDM (with robust standard errors)

	Model 1		Model 2	
Fixed Effect	Coefficient	S.E.	Coefficient	S.E.
Intercept	0.918***	0.006	0.931***	0.005
principal gender			-0.018*	0.008
principal experience			0.000	0.001
school level-elementary			-0.062***	0.017
school level-secondary			0.044**	0.016
school enrollment			0.004***	0.001
percent of minority student			0.000	0.000
percent of NSLP student			0.000	0.000
school location-urban			0.012	0.017
school location-suburban			0.001	0.013
principal data sources			0.023*	0.012
school performance			-0.018*	0.007
Random Effect	Variance	*p*	Variance	*p*
Level-2 between district	0.04152	0	0.04050	0
Level-1 within district	0.01118		0.01095	
Intra-class correlation	0.787855787		0.787172012	

Note: *** p < .001; ** .001 < p < .01; * .01 < p < .05

Effects of Composite DDIDM on SDIDM

The results of Model 3, Model 4, and Model 5 were included in Table 5. The results of Model 3 revealed that two district characteristics: district total enrollment and percentage of students enrolled in the national free and reduced lunch program were not significant. Comparing with Model 2, the improvement of Model 3 at level-2 could be gauged by the proportion of variance explained: (0.0405 – 0.0405) / 0.0405= 0. This indicated that adding the two district control variables could hardly reduce level-2 variance.

The results of Model 4 showed the focused composite variable—DDIDM was significant and positive (β= 0. 152, p < .0001, effect size = 7.66). The improvement of Model 4 at level-2 was: (0. 0405– 0.0163) / 0. 0405= 0. 5968. This indicated that after including DDIDM, level-2 variance was reduced by 59.68%. The results suggested that there was a strong, statistically significant and positive relationship between district data-informed decision-making and school data-informed decision-making.

Effects of Specific DDIDM on Specific SDIDM

The results of Model 5 showed that, school's decision making were significantly and positively influenced by all five district-level decision making areas: including district's application of student performance report for evaluating the progress of students (β= 0. 138, p < .0001, effect size = 7.26); determining the next year's instructional focus (β= 0. 081, p < .01, effect size = 4.28); realigning the curriculum with assessment and/or other indicator criteria (β= 0. 082, p < .05, effect size = 4.3); informing parents and the community of the district's and/or school's progress (β= 0. 406, p < .0001, effect size = 21.37); and prompting school-level initiatives for improvements (β= 0. 102, p < .0001, effect size = 5.35). The improvement of

Model 5 over Model 3 at level-2 was gauged by the proportion of variance explained: (0. 0405–0. 0152) / 0. 0405= 0. 6257. This indicated that after including the five district applications of student performance for decision making, level-2 variance was substantially reduced by 62.57%.

Table 6

Results of the Impact of DDIDM on SDIDM (with robust standard errors)

Fixed Effect	Model 3		Model 4		Model 5	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Intercept	0.932***	0.005	0.933***	0.003	0.934	0.003
principal gender	-0.018*	0.008	-0.005	0.006	-0.004	0.005
principal experience	0.000	0.001	0.000	0.000	0.000	0.000
school level-elementary	-0.061***	0.017	-0.026**	0.010	-0.026**	0.010
school level-secondary	0.043**	0.016	0.003	0.010	0.002	0.010
school enrollment	0.004***	0.001	0.003***	0.001	0.002***	0.001
percent of minority student	0.000	0.000	0.000	0.000	0.000	0.000
percent of NSLP student	0.000	0.000	0.000	0.000	0.000	0.000
school location-urban	0.015	0.018	0.010	0.011	0.011	0.011
school location-suburban	0.003	0.015	0.003	0.008	0.003	0.008
principal data sources	0.023	0.012	0.015	0.009	0.012	0.010
school performance	-0.018*	0.007	-0.009	0.006	-0.010	0.006
level-2						
percent of NSLP in district	0.000	0.000	0.000	0.000	0.000	0.000
district enrollment	0.000	0.000	0.000	0.000	0.000	0.000
DDIDM			0.152***	0.007		
District-level data-informed decision areas						
Evaluating students					0.138***	0.041
Determining focus					0.081**	0.029
Realigning curriculum					0.082*	0.033
Informing parents					0.406***	0.054
Prompting initiatives					0.102***	0.023
Random effects	Variance	*p*	Variance	*p*	Variance	*p*
Level-2	0.0405	0	0.0163	0	0.0152	0
Level-1	0.011		0.0113		0.0114	
Conditional ICC	0.7869		0.5907		0.5712	
Proportion of variance explained at level-2	Over Model 2		Over Model 3		Over Model 4	
	0.0000		0.5968		0.6257	

Note: *** p < .001; ** .001 < p < .01; * .01 < p < .05

Lastly, a few school- and district-level background variables were found to be statistically significant. At school level, smaller, elementary schools school (vs middle and high school) were associated with lower level of school data-informed decision-making.

Findings of Study 3

Variance of DDIDM

After conducting the analysis for the unconditional model, the results were presented in Table 7. The results showed that the level-2 variance was 0.7165. Therefore, the intra-class correlation (ICC) was 0.7165/ (0.7165+3.29) = 0.1788, suggesting that adding a level-2 model can potentially explain 17.9% of the outcome variance.

The results of two control models (Model 2 and Model 3) were also included in Table 7. The conditional ICC is 0.1665 and 0.1647, respectively, indicating that after controlling for school level background, 16.65% of the outcome variability existing between districts; after controlling for district level background, 16.47% of the outcome variability existing between districts.

Table 7

Results of the Unconditional Model and the Two Control Models on School Performance (Population Average with robust standard errors)

Fixed Effect	Model1		Model2		Model3	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Intercept	0.1563**	0.051	0.022	0.048	0.027	0.049
Level 1 School level						
principal gender			0.175	0.105	0.176	0.105
principal experience			0.014	0.007	0.014*	0.007
school level-elementary			-0.261	0.167	-0.241	0.168
school level-secondary			-0.005	0.171	-0.005	0.171
school enrollment			-0.026*	0.012	-0.026*	0.012
percent of minority student			-0.005	0.002	-0.004	0.002
percent of NSLP student			-0.013***	0.003	-0.009***	0.003
school location-urban			0.052	0.154	0.090	0.158
school location-suburban			-0.042	0.117	-0.002	0.119
principal data sources			0.099	0.156	0.103	0.156
average teacher experience			0.002	0.009	0.002	0.008
percent of teachers teaching to high standard			0.019***	0.003	0.019***	0.003
percentage of regular/probational teacher certificate			0.269	0.362	0.274	0.364
percent of male teachers			-0.509*	0.213	-0.506*	0.211
Level 2 District level						
percent of NSLP in district					-0.006*	0.003
district enrollment					0.000	0.000

Table 7—Continued

Random effects	Variance	p	Variance	p	Variance	p
Variance of fitted values	0.1563		0.5090		0.4170	
Level-2	0.7165	0	0.6571	0	0.6487	0
Level-1	3.29		3.29		3.29	
Conditional ICC	0.1788		0.1665		0.1647	
Proportion of variance explained at level-2	0.0376		0.1142		0.0957	

Note: *** $p < .001$; ** $.001 < p < .01$; * $.01 < p < .05$

Effects of Specific TDIDM, SDIDM, DDIDM on Performance

The results of Model 4, Model 5, and Model 6 were included in Table 8. The results of Model 4 suggested that none of the school average teacher data-informed decision-making were significant. Although the results of Model 5 showed that none of the five focused school data-informed decision-making variables were significant, the results of Model 6 suggested that using student performance report to evaluate student progress at school level was significant and positive after adding the district data-informed decision-making predictors (β= 0. 640, $p < .05$), which means that expected odds of passing all standards of schools that use performance report to evaluate student progress are exp (0.640) =1.896 times the odds of passing all standards of schools that do not use performance report to evaluate student progress holding constant random effect. In other words, the probability of passing all standards differs for schools which do and do not use student performance report to evaluate student progress by 65.5%. The results of Model 6 also showed that none of the five district data-informed decision-making variables were significant.

Table 8

Results of the Impact of TDIDM, SDIDM, and DDIDM on School Performance (Population Average with robust standard errors)

Fixed Effect	Model4		Model5		Model6	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Intercept	0.029	0.049	0.034	0.049	0.035	0.049
Level 1 School level						
principal gender	0.190	0.105	0.193	0.105	0.189	0.105
principal experience	0.014*	0.007	0.014*	0.007	0.014*	0.007
school level-elementary	-0.224	0.169	-0.218	0.168	-0.223	0.168
school level-secondary	0.005	0.171	-0.0015	0.169	0.004	0.170
school enrollment	-0.025*	0.012	-0.024*	0.012	-0.024*	0.012
percent of minority student	-0.005	0.002	-0.005	0.002	-0.004	0.002
percent of NSLP student	-0.009***	0.003	-0.009***	0.003	-0.009***	0.003
school location-urban	0.099	0.158	0.116	0.157	0.132	0.159
school location-suburban	0.001	0.120	-0.002	0.119	0.005	0.120
principal data sources	0.104	0.158	0.107	0.157	0.112	0.158
average teacher experience	0.003	0.009	0.003	0.008	0.003	0.009
percent of teachers teaching to high standard	0.019***	0.003	0.019***	0.003	0.018***	0.003
percentage of regular/probation teacher certificate	0.265	0.363	0.243	0.359	0.246	0.358
percent of male teachers	-0.479*	0.213	-0.473*	0.212	-0.468*	0.212
Average teacher-level data-informed decision areas						
Group students	0.097	0.079	0.092	0.079	0.093	0.080
Assess areas	0.030	0.121	0.046	0.120	0.040	0.120
Adjust curriculum	-0.040	0.126	-0.046	0.125	-0.041	0.126
School-level data-informed decision areas						
Evaluating students			0.556	0.302	0.640*	0.306
Determining focus			-0.223	0.218	-0.250	0.220
Realigning curriculum			-0.349	0.195	-0.287	0.194
Informing parents			0.132	0.299	0.222	0.298
Prompting initiatives			-0.192	0.221	-0.151	0.227

Table 8—Continued

Level 2 District level						
percent of NSLP student in district	-0.006*	0.003	-0.006*	0.003	-0.007*	0.003
district enrollment	0.000	0.000	0.000	0.000	0.000	0.000
District-level data-informed decision areas						
Evaluating students					-0.359	0.293
Determining focus					0.146	0.253
Realigning curriculum					-0.480	0.274
Informing parents					0.192	0.446
Prompting initiatives					0.035	0.287
Random effects	Variance	p	Variance	p	Variance	p
Variance of fitted values	0.423		0.44		0.434	
Level-2	0.6463	0	0.6395	0	0.6427	0
Level-1	3.29		3.29		3.29	
Conditional ICC	0.1642		0.1627		0.1634	
Proportion of variance explained at level-2	0.097		0.1007		0.0994	

Note: *** p < .001; ** .001 < p < .01; * .01 < p < .05;

Some district and school-level background variables were found to be statistically significant. The odds of a school passing all state and district standards would increase if the school had a more experienced principal, smaller school size, fewer students on free and reduced-price lunch, more high-standard teachers, and more female teachers. At the district level, a school in a district that had fewer students on free and reduced-price lunch were also more likely to pass all standards.

Summary

In this chapter, I presented detailed statistical findings of each of the studies. In the first study, statistically significant associations were identified, however the effect sizes were small, and a very small proportion of variance was explained. In the second study statistically significant associations were identified with large effect sizes, and a large proportion of variance was explained. In the third study, only one school level data-informed decision-making area was statistically significant. In the following chapter, I will present a discussion on the implication of my findings.

CHAPTER V

DISCUSSION

Viewing "using data" as a coupling mechanism and "decision-making" as a coupling element, I found that the K-12 system appears to be "loosely coupled" between the school and classroom levels, but "tightly coupled" between the district and school levels. I also found that most data-informed decision-making of districts, schools, and teachers were not significantly associated with school performance. My findings have implications for educational organizational theory and school improvement.

Discussion of Major Findings

The present three studies reveal some interesting relationship within the education system by viewing "data using" as a coupling mechanism and "decision-making" as a coupling element.

Tightly Coupled District/School but Loosely Coupled School/Classroom

My first study focuses on the relationship between the school-level decision making based on student performance data and teacher-level instructional decision making based on student's state or district achievement test data. The results indicated that, across all five school-level decision making areas, only one area—determining the next year's instructional focus—had statistically significant influence. Nevertheless, such statistically significant influence with small effect size is too weak to support a conclusion that data has changed the loose coupling status of schools in the age of standards-based reform and accountability movement. The schools and classrooms were largely loosely coupled in terms of data-informed decision-making when

the national survey was administered in 2003-04. The technical core, i.e., teachers' data-based instructional practice, had little or no association with the school-level decision making.

My second study examined the relationship between the district-level decision making based on student performance data and school-level decision making based on student performance data. It was found all five specific data-informed district decision makings were significantly associated with the extent of school data-informed decision-making. Especially, district application of student performance reports for evaluating the progress of students, informing parents and the community of the district's and/or school's progress, and prompting school-level initiatives for improvements are the top three of the five areas in terms of effect size. Therefore, it could be concluded that the districts and schools were tightly coupled in terms of the data-informed decision-making in the 2003-04 school year.

Based on the findings of both studies, I argue that even in the age of standardization and accountability, teachers' technical core is still hard to change. One may argue that it is not surprising to see that teacher decision makings are not highly correlated with school decision-making, while school decision makings are highly correlated with district decision-making. Although the teacher data-informed decision-making variables are quite different from school data-informed decision-making variables in study 1, and school data-informed decision-making variables and district data-informed decision-making variables are almost the same in study 2. It is true that types of decision makings vary across different levels and various positions of educational system. As Conley (1991) noted "administrators make strategic decisions outside of classrooms and teachers make operational decisions inside of classrooms" (pp. 237-238). However, it is important to note that (a) the school and district data-informed decision-making variables from the SASS are similar to the teacher data-informed decision-making variables in

that both groups of variables emphasized decision making regarding instruction and curriculum rather than management; and (b) the effect sizes from both studies show that if a district employs data-informed decision-making extensively, its schools are highly likely to employ data-informed decision-making extensively. However, a school's effort to use data to inform decisions is not associated with teacher's extent of using data to make informed decisions.

Accountability Oriented Data-informed Decision-making

While data-informed decision-makings are tightly coupled between districts and schools, and are loosely coupled between schools and classrooms, my findings from study 3 suggested that only one specific data-informed decision-making type—schools using student performance data to evaluate student progress—can significantly improve school performance. Other types of decision-making still can hardly improve school performance.

Based on the literature on standardization, accountability, and organizational learning, each specific data-informed decisions by districts, schools, and teachers can be categorized into these three major types—accountability-oriented, standardization-oriented, and organizational learning oriented. For example, at the teacher level, "assessing areas where the teacher needs to strengthen his or her content knowledge or teaching practices" is obviously an organizational learning oriented decision-making. At the school and district levels, "realigning the curriculum, such as with content standards and/or other indicator criteria", is a typical standardization oriented decision-making, while "evaluating the progress of students in this district or schools" is a typical accountability oriented decision-making, and "determining the next year's instructional focus" and "prompting school-level initiatives for improvement" sound more of organizational learning oriented. Additionally, "informing parents and the community of the district's and/or

school's progress" could work as both accountability oriented and organizational learning oriented.

Leithwood et al. (2004) argued that "the chance of any reform improving student learning is remote unless district and school leaders agree with its purposes and appreciate what is required to make it work" (p. 7). It is noteworthy that previous studies found that the patterns of data use were weakly related to student achievement in elementary schools (Anderson, Leithwood, & Strauss, 2010). Given the fact that the accountability oriented data-informed decision-making—schools using student performance data to evaluate student progress—is the among dominant schemes of reform that received great attention by both schools and districts, it is not surprising to see that schools using student performance data to evaluate student progress is the only type of data-informed decision-making that can effectively improve student performance.

Implications

Against the backdrop of standards-based reform and accountability movement in the last 15 years, which focus on tightening the system as the way to improve K-12 schools, I conducted three related studies to empirically test the "loosely coupled" theory with a nationally representative data in the US. My findings suggest that the widespread blanket statement of "schools as loosely coupled organizations" is not accurate. Thus, the dominant reform agenda in the last 15 years based on this blanket statement is called into question.

Blanket Statement of "Schools as Loosely Coupled Organizations"

Although there is a great deal of literature about the looseness of educational organizations, many educational organization researchers and policy makers used "schools as

loosely coupled organizations" as a blanket statement. This statement is inaccurate unless it has been clarified as to what kind of relationship within education systems was discussed.

Clarification of levels and domains. It is very important for organizational theorists to understand the multi-layer, complexity, and dynamics of loose coupling of educational organization. My studies add new evidence to the classic argument on whether the degree of looseness varied within organizations (see Ingersoll, 1993; 1994; 1996; Shen & Ma, 2006). The findings of my studies confirmed that the degree of looseness varied not only by the organizational levels, but also by the domains – data-informed decision-makings are found to be tightly coupled between districts and schools, but loosely coupled between schools and classrooms. At the same level, there is also a large disparity in terms of the strength of coupling among each specific data-informed decision-making area. Therefore, it may be arbitrary to state that the education system is loosely coupled without identifying which domains and levels are specifically discussed.

Clarification of coupling elements and coupling mechanism. Clarification of organizational levels and domains is not enough. It is also critical to define the coupling elements and their coupling mechanisms in loose coupling studies (Rowan, 2002). School elements are loosely coupled or tightly coupled under varied effects of different mechanism. H. D. Meyer (2002a) insightfully pointed out that the discussion of loose coupling in the literature frequently misses "any kind of linkage between any kind of organizational 'element', from schedules, routines, or policies, to attitudes, norms, and preferences" (p. 536). In this paper, the existing knowledge (i.e., the introduction of data into education system by federal and state governments has changed the ambiguity of decision making) informed me that there could be a new coupling mechanism (i.e., using data) for a coupling element (i.e., decision making at different levels).

Without a clarification of coupling element and mechanism, the judgment on loose coupling of organizations is not well advised.

Better Education through Collaborations across the System

The inaccurate blanket statement of "schools as loosely coupled organizations" directly leads to the domination of the top-down managerialism scheme. In the last 15 years, the U.S. education has witnessed efforts to tighten the education system via curriculum standards, accountability tests, and evaluation.

There is no doubt that in the last 15 years, the system is becoming more and more formal, rationalized, and structured, especially at the district and school level. My findings from study 2 support the conclusion that schools and districts are coherent in decision making areas, including instruction and curriculum, although many earlier studies suggested the opposite (e.g., Crowson & Morris, 1985; Deal and Celotti, 1980). However, it is questionable if the top-down managerialism scheme is useful across all levels of education system. My findings from the first study agree with many early and recent studies (e.g., Shen & Ma, 2006) that the classroom teachers are still decoupled from schools.

Turn focus to isolated teachers: organizational learning. A large body of work on loose coupling theory highlights two sources of loose coupling at teacher level–external inspection as ritual rather than control (Deal & Celotti,1980; J. W. Meyer & Rowan, 1978; Weick, 1976) and isolation of teachers (Goodlad & Klein,1975; J. W. Meyer & Rowan, 1978; Weick, 1976). The top-down managerialism scheme obviously is able to solve the first problem, however, at the same time, may have already made the second problem even worse. Notably at the teacher level, studies have documented evidence that teachers are becoming more isolated

from their principals and colleagues (Grissom et al., 2014; Hargreaves & Goodson, 2006; Sheldon & Biddle, 1998).

Instead of the top-down managerialism scheme, it is time for policy makers to consider another scheme to better involve and motivate teachers to achieve the goal of bringing better education to children—the bottom-up, organizational learning scheme. The organizational learning scheme, which operates under the assumption of working with, rather than against, loose-coupling, has been neglected in the policy agenda over the past decades of educational reform.

The organizational learning scheme is compatible with loose coupled nature of educational organizations in many ways. Loose coupling organizations attach great importance to trust relationship (Rowan & Meyer, 1976), emphasize cultural linkages rather than traditional bureaucratic linkages (Firestone & Wilson, 1985), and focus on production of knowledge (Sanchez & Mahoney, 1996). These attributions appear to be consistent with the model of organizational learning, which advocates professional autonomy, mutually supportive relationships, shared norms, acquisition of knowledge and skills, and collective purpose (Stoll et al., 2006).

Culture for supporting. Organizational learning scheme highlighted the importance of a culture of supporting teachers. The top-down, new managerialism strategy assumes that in order to change the technical core of teaching, the loosely-coupled organization has to be tightened. Such an assumption only emphasizes the bureaucratic facet of a school, but misses its cultural aspect. Culture is always the centerpiece of school reform (Sarason, 1982). Compelling evidence documented by Bryk et al. (2009) from Chicago public school reform showed that high trust was associated with stronger developments across organizations. Firestone and Wilson (1985) noted

that school leaders influence the technical core through both the bureaucratic and cultural linkages, while bureaucratic linkages create restraints and opportunities for teachers, and cultural linkages shape teachers' vision.

Particularly, a learning organization advocates the importance of a culture of collaboration, which engages educators in "an ongoing cycle of questions that promote deep team learning" (DuFour, 2004, p. 9). Nelson and Slavit (2007) highlighted the sense of professional trust when teachers work collaboratively with each other in using data to improve their practice. Existing literature has documented evidence of a changed technical core after schools adopted professional learning community models (Dunne, Nave, & Lewis, 2000; Harris & Jones, 2010; E. R. Hollins, McIntyre, DeBose, K. S. Hollins, & Towner, 2004; Louis & Marks, 1998; Strahan, 2003). It noteworthy that the recent education initiatives, the Common Core State Standard Initiative (CCSSI) were criticized for its heavy reliance on the extensive teacher performance evaluation based on student performance, while paying little attention to the learning culture (Mathis, 2010).

A more complicated supporting system for decision making. In addition to the cultural development, advanced technologies also help bring the ideas of learning organization into reality. The new technologies will improve teacher capacity of using data and allow data-informed decision-makings to be seamlessly integrated into the learning organization, which will change the isolation of teachers in a deep way. Pioneer efforts on creating Networked Improvement Communities by the Carnegie Foundation suggested embedding data into everyday work to ensure continuous improvement (Bryk, Gomez, & Grunow, 2011). The recent developments of learning analytic tools are making "big data" in classrooms possible. Teachers will be able to perform data analysis, and design their learning objectives as part of their daily

work (Dyckhoff et al., 2012; L. Johnson et al., 2011). These new data-informed approaches to educational improvement indicate that organizational learning oriented data-informed decision-making may be an important tool of change, supported by a loose-coupling school context at the teacher level. By allowing teachers to constantly improve their practice through utilizing data independently or cooperatively, the merits of loose coupling, such as independency, adaptability, diversity, and self-efficacy, could be maintained, while the demerits of loose coupling, such as ambiguity and obstinacy, could be minimized.

Limitations

One of the purposes of my studies was to test the "school as a loosely coupled organization" theory through examining the association between school-level data-informed decision-making and teacher-level instructional data-informed decision-making, as well as the association between district-level data-informed decision-making and school-level data-informed decision-making. To examine the postulate, a series of nested models were applied to nationally representative teacher, principal/school, and district samples. The two-level HLM model was used as the primary statistical method to build these models. My studies had strengths in (a) addressing the traditional depiction of the educational organizations as a loosely coupled system and current advocacy of data-informed decision-making, (b) using a nationally representative sample with a rich array of variables at the teacher, principal/school, and district levels, and (c) employing an appropriate statistical method that capitalized the data structure.

However, the present study also has its limitation and points out directions for future efforts. The first limitation is related to the timeliness of the data sets used in the studies. My studies used SASS2003-04 data, which was collected more than ten years ago. However, the more recent data from SASS did not include the variables I need for the studies and thus do not

fit the purpose of this dissertation paper. I admit that the data was collected more than 10 years ago, and it might be "outdated" for most studies. Nevertheless, I would like to justify my use of the 2003-2004 SASS data with the following reasons: First, although "loose-coupling" school system has been a classic topic for more than two decades, relatively few studies investigated the topic with large-scale empirical data (Rowan, 2002). To our knowledge, SASS 2003-04 is the only national-level survey that provides items related to educators' data-informed decision-making activity at both classroom level and school level. Unfortunately, the following cycle of SASS surveys, SASS 2007-08 and SASS 2011-12, do not have similar items. Second, although SASS 2003-04 is a relatively old data set, it was administered just after a very important time point of the U.S. history of education – the NCLB in 2002. I feel it is still meaningful to capture the picture of how data-informed decision-makings at the teacher, school and district levels interacted at the beginning point of NCLB.

The second limitation of the present study is related to first limitation. My data sets only have one time point to study, thus I can only use cross-sectional modeling instead of longitudinal modeling to estimate the influence of data-informed decision-making on school performance. It is difficult to tell causal relationship between student outcome and various organizational variables with cross-sectional data, as "the nature of life in schools is both complex and cyclical" (Hallinger & Heck, 2011, p. 153). Louis, Leithwood, Wahlstrom and Anderson (2010) suggested that "longitudinal evidence that displays trends and trajectories has greater potential than snapshot data for informing educators' school-improvement activity" (pp. 184-185). I believe it would be more interesting to evaluate the relationship of data-informed decision-making at three edu cational levels and their influence on student achievements over the years.

The third limitation of my study is that I used quantitative data only. My recommendation for future studies is that it would be interesting to continue this line of research by combining empirical evidence from both a qualitative approach and a quantitative approach. While the quantitative approach can help establish relationships and offer more generalizable conclusions, the lens of a qualitative approach may deepen our insights of the relationships and effects of data-informed decision-making at different levels by gathering more information and successful examples.

REFERENCES

Amrein, A. L., & Berliner, D. C. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives, 10*(18). Retrieved from http://epaa.asa.edu/epaa/v10n18/

Amrein, A., & Berliner, D. C. (2003). Re-analysis of NAEP math and reading scores in states with and without high-stakes tests: Response to Rosenshine. *Education Policy Analysis Archives*, *25* (11). Retrieved from http://epaa.asu.edu/epaa/v11n25/

Anderson, S., & Kumari, R. (2009). Continuous improvement in schools: Understanding the practice. *International journal of educational development*,*29*(3), 281-292.

Anderson, S., Leithwood, K., & Strauss, T. (2010). Leading Data Use in Schools: Organizational Conditions and Practices at the School and District Levels. *Leadership and Policy in Schools, 9*(3), 292–327. http://doi.org/10.1080/15700761003731492

Babbie, E. (2010). *The practice of social research* (12th Ed.). Belmont, CA: Thomson Wadsworth.

Ball, S. J. (1998). Big Policies/Small World: An introduction to international perspectives in education policy. *Comparative Education*, *34*(2), 119–130. http://doi.org/10.1080/03050069828225

Ball, S. J. (1997). Policy sociology and critical social research: A personal review of recent education policy and policy research. *British educational research journal, 23*(3), 257-274.

Ballou, D., & Springer, M. G. (2008). *Achievement trade-offs and No Child Left Behind (Working paper).* Peabody College of Vanderbilt University. Nashville, TN.

Bernhardt, V. L. (2013). *Data analysis for continuous school improvement.* New York, NY: Routledge.

Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas accountability system. *American Education Research Journal, 42* (2), 231-268.

Bottery, M. (1996). The challenge to professionals from the new public management: Implications for the teaching profession. *Oxford Review of Education*, *22*(2), 179–197. http://doi.org/10.1080/0305498960220206

Bracey, G. W. (2008). Assessing NCLB. *Phi Delta Kappan*, *89*(10), 781-782.

Braun, H. (2004). Reconsidering the Impact of High-Stakes Testing. *Education Policy Analysis Archives*, *12*(1).

Bridge, C. A., Compton-Hall, M., & Cantrell, S. C. (1997). Classroom writing practices revisited: The effects of statewide reform on writing instruction. *Elementary School Journal, 98*(2), 151–170.

Bryk, A. S., Sebring, P. B., Allensworth, E., Luppescu, S., & Easton, J. (2009). *Organizing schools for improvement: Lessons from Chicago*. Chicago, IL: University of Chicago Press.

Bryk, A. S., Gomez, L. M., and Grunow, A. (2011). Getting ideas into action: Building networked improvement communities in education. In M.T. Hallinan (Ed.), *Frontiers in Sociology of Education* (pp. 127-162). New York, NY: Springer.

Burian-Fitzgerald, M., McGrath, D. J., & Plisko, V. (2003). Mathematics teachers' familiarity with standards and their instructional practices: 1995 and 1999. *Education Statistics Quarterly, 5*(1), 53–56.

Bush, T. (2003). *Theories of Educational Leadership and Management*. London: Sage Publishing

Butty, J. L. M. (2001). Teacher instruction, student attitudes, and mathematics performance among 10th and 12th grade black and Hispanic students. *Journal of Negro Education, 70*(1–2), 19–37.

Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Educational Evaluation and Policy Analysis, 24*(4), 305-331. http://doi.org/10.3102/01623737024004305

Celio, M.B., & Harvey, J. (2005). *Buried treasure: Developing a management guide from mountains of school data.* Seattle, WA: Center on Reinventing Public Education.

Chance, P. L., & Anderson, R. B. (2003, April). *The principal's role in standards-based reform: Linking accountability to instructional improvement*. Paper presented at the Annual meeting of the American Educational Research Association, Chicago, IL. (ERIC Document Reproduction Service No. ED478258)

Chatterji, M. (2002). Models and methods for examining standards-based reforms and accountability initiatives: Have the tools of inquiry answered pressing questions on improving schools?. *Review of Educational Research*,*72*(3), 345-386.

Cheong, Y. F., Fotiu, R. P., & Raudenbush, R. W. (2001). Efficiency and robustness of alternative estimators for 2- and 3-level models: The case of NAEP. *Journal of Educational and Behavioral Statistics*, 26, 411-429.

Chrispeels, J. H., Burke, P. H., Johnson, P., & Daly, A. J. (2008). Aligning Mental Models of District and School Leadership Teams for Reform Coherence. *Education and Urban Society*, *40*(6), 730–750.

Chrispeels, J. H., Castillo, S., & Brown, J. (2000). School leadership teams: A process model of team development. School Effectiveness and School Improvement, 11(1), 20-56.

Chudowsky, N., & Chudowsky, V., & Kober, N. (2007). *Answering the question that matters most: Has student achievement increased since No Child Left Behind*? Washington, DC: Center on Educational Policy.

Clarke, M., Shore, A., Rhoades, K., Abrams, L., Miao, J., & Li, J. (2003). *Perceived effects of state-mandated testing programs on teaching and learning: Findings from interviews with educators in low-, medium-, and high-stakes states.* Chestnut Hill, MA: National Board on Educational Testing and Public Policy.

Clune, W. (1998). *Toward a theory of systemic reform: The case of nine NSF statewide systemic initiatives*. (National Institute for Science Education Research Monograph No. 16). Madison, WI: National Institute for Science Education Research.

Coburn, C. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher, 32*(6), 3-12.

Coburn, C. E. & Turner, E. O. (2012). The practice of data use: An introduction. *American Journal of Education*, *118*(2), 99-111.

Coburn, C. E., & Talbert, J. E. (2006). Conceptions of evidence use in school districts: Mapping the terrain. *American Journal of Education*, *112*(4), 469-495.

Coburn, C. E. (2001). Collective sense-making about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis*.

Codd, J. (2005). Teachers as 'managed professionals' in the global education industry: The New Zealand experience. *Educational Review*, *57*(2), 193-206.

Cohen, D. (1995). What standards for national standards?. *Phi Delta Kappan*,*76*(10), 751.

Cohen, D. (1990). A revolution in one classroom: The case of Mrs. Oublier. *Educational Evaluation and Policy Analysis*, *12*(3), 311-329.

Cohen, M., March, J., & Olsen, J. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly*, *17*(1), 1-25.

Conley, S. (1991). Review of research on teacher participation in school decision making. *Review of Research in Education, 17*, 225-266.

Cosner, S. (2011). Teacher learning, instructional considerations and principal communication: Lessons from a longitudinal study of collaborative data use by teachers. *Educational Management Administration and Leadership, 39*(5), 568–589.

Crowson, R. L. & Morris, V. C. (1985). Administrative control in large-city school systems: An investigation in Chicago. *Educational Administration Quarterly, 21*(4), 51-70.

Cuban, L. (1990). Reforming again, again, and again. *Educational researcher, 19*(1), 3-13.

Daly, A. J. (2012). Data, dyads, and dynamics: Exploring data use and social networks in educational improvement. *Teachers College Record*, *114*(11), 1-38.

Darling-Hammond, L., & McLaughlin, M. W. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappan, 76*(8), 597

Datnow, A., & Hubbard, L. (2015). Teacher capacity for and beliefs about data-driven decision making: A literature review of international research. *Journal of Educational Change*, *17*(1), 7–28.

Datnow, A., Park, V., & Kennedy-Lewis, B. (2013). Affordances and constraints in the context of teacher collaboration for the purpose of data use. *Journal of Educational Administration, 51*(3), 341–362.

Datnow, A., Park, V., & Wohlstetter, P. (2007). *Achieving with data: How high performing school systems use data to improve instruction for elementary students*. Los Angeles, CA: Center

on Educational Governance, Rossier School of Education, University of Southern California.

Deal, T. E., & Celotti, L. D. (1980). How much influence do (and can) educational administrators have on classrooms? *Phi Delta Kappan*, *61*(1), 471-473.

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

Dee, T. S., Jacob, B. A., Hoxby, C. M., & Ladd, H. F. (2010). The impact of No Child Left Behind on students, teachers, and schools. *Brookings Papers on* Economic Activity. 149-194.

Deming, W. E. (1986). *Out of the Crises.* MIT Center for Advanced Engineering Study, Cambridge, MA.

Dennis, D.V. (2008). Are assessment data really driving middle school reading instruction? What we can learn from one student's experience. Journal of Adolescent & Adult Literacy, 51(7), 578-587.

Desimone, L. M. (2013). Reform before NCLB: Standards-Based Reform before NCLB Was More Successful Because It Sought Teacher Buy-In and Was Less Punitive. *Phi Delta Kappan*, *94*(8), 59.

Diamond, J. B., & Cooper, K. (2007). The uses of testing data in urban elementary schools: Some lessons from Chicago. *Yearbook of the National Society for the Study of Education, 106*(1), 241–63.

DiMaggio, P. J., & Powell, W. W. (Eds.). (1991). *The new institutionalism in organizational analysis* (Vol. 17). Chicago, IL: University of Chicago Press.

DuFour, R. (2004). What is a" professional learning community*?". Educational Leadership, 61*(8), 6-11.

Dunn, K. E., Airola, D. T., Lo, W. J., & Garrison, M. (2013). What teachers think about what they can do with data: Development and validation of the data driven decision-making efficacy and anxiety inventory. *Contemporary Educational Psychology*, *38*(1), 87-98.

Dunne, F., Nave, B., & Lewis, A. (2000). Critical friends groups: Teachers helping teachers to improve student learning. *Phi Delta Kappan*, *28*(4), 31-37.

Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A., & Schroeder, U. (2012). Design and Implementation of a Learning Analytics Toolkit for Teachers. *Educational Technology & Society, 15*(3), 58–76.

Elmore, R. F. (2000). *Building a new structure for school leadership.* Washington, DC: Albert Shanker Institute.

Elmore, R.F. (1993). The role of local school districts in instructional improvement. In S. Fuhrman, (Ed.), *Designing coherent education policy: Improving the system* (pp. 96-124). San Francisco, CA: Jossey-Bass Publishers.

Erickson, D. K., & Niess, M. L. (1996). Focusing on NCTM's standards: Teachers' choice and decisions related to students achievement in middle school mathematics. *Research in Middle Level Education Quarterly, 19*(2), 23–42.

Falk, J., & Drayton, B. (2004). State testing and inquiry-based science: Are they complementary or competing reforms? *Journal of Educational Change, 5*(4), 345–387.

Farkas, S., Johnson, J., & Duffett, A. (2003). *Rolling up their sleeves: Superintendents and principals talk about what's needed to fix public schools*. New York, NY: Public Agenda.

Figlio, D. N., Rouse, C. E., & Schlosser, A. (2009). *Leaving no child behind: Two paths to school accountability* (Working paper). Northwestern University Institute for Policy Research, Chicago, IL.

Firestone, W. (2009). Accountability nudges districts into changes in culture. *Phi Delta Kappan*, *90*, 670-676.

Firestone, W. A., & González, R. A. (2007). Culture and processes affecting data use in schools. In P. A. Moss (Ed.), *Evidence and decision making* (pp. 132-154). Malden, MA: Blackwell

Firestone, W. A., & Wilson, B. L. (1985) Using bureaucratic and cultural linkages to improve instruction: The principal's contribution, *Educational Administration Quarterly*, *21*(2), 7-30.

Firestone, W., Mayrowetz, D., and 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.

Fitchett, P. G., Heafner, T. L., & Lambert, R. (2014). Examining elementary social studies marginalization: A multilevel model. *Educational Policy, 28*(1), 40-68.

Fitzgerald, J. H. (1993). Management practices: A profile of district-level supervisory activity in one school district. *Journal of Curriculum and Supervision, 8*(2), 128-39.

Fuchs, L. S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review*, *33*(2), 188–192.

Fullan, M. (2007). Educational reform as continuous improvement. In Hawley, W. (Ed.), *The Keys to Effective Schools: Educational Reform as Continuous Improvement* (pp. 1–12). Corwin, CA: Thousand Oaks.

Fusarelli, L. D. (2002). Tightly coupled policy in loosely coupled systems: Institutional capacity and organizational change. *Journal of Educational Administration*, *40*(6), 561-575.

Goldring, E. B. & Hallinger, P. (1992). *District control contexts and school organizational processes*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Goldspink, C. (2007). Rethinking educational reform: a loosely coupled and complex systems perspective. *Educational Management Administration & Leadership*, *35*(1), 27-50.

Goodlad, J. I. (1994). *What schools are for (2nd ed.)*. Bloomington, IN: Phi Delta Kappa Educational Foundation.

Goodlad, J., Klein, M.F. & Associates. (1975). *Looking behind the classroom door.* Worthington, OH: Charles A. Jones Publishing Company.

Grant, S. G. (2001). An uncertain lever: Exploring the influence of state-level testing in New York State on teaching social studies. *Teachers College Record, 103*(3), 398–426.

Grant, T. J., & Kline, K. (2000). *Understanding teachers' changing beliefs & practice while implementing a reform curriculum.* Paper presented at the Annual meeting of the American Educational Research Association, New Orleans, LA.

Griffiths, D. E., Hart, A. W., & Blair, B. G. (1991). Still another approach to administration: Chaos theory. *Educational Administration Quarterly*, *27*(3), 430-451.

Grissom, J. A., Nicholson-Crotty, S., & Harrington, J. R. (2014) Estimating the effects of no child left behind on teachers' work environments and job attitudes. *Educational Evaluation and Policy Analysis, 20*(10), 1-20.

Gummer, E., & Mandinach, E. (2015). Building a conceptual framework for data literacy. Teachers College Record, 117(4), 1–22.

Gunter, H. M., & Fitzgerald, T. (2013). New Public Management and the modernisation of education systems 1. *Journal of Educational Administration and History*, *45*(3), 213–219.

Hallinger, P., & Heck, R. H. (2011). Conceptual and methodological issues in studying school leadership effects as a reciprocal process. *School Effectiveness and School Improvement, 22*(2), 149-173.

Halverson, R., Grigg, J., Prichett, R., & Thomas, C. (2007). The new instructional leadership: Creating data-driven instructional systems in school. *Journal of School Leadership, 17*(2), 159-194.

Hamilton, L. S., McCaffrey, D. F., Stecher, B. M., Klein, S. P., Robyn, A., & Bugliari, D. (2003). Studying large-scale reforms of instructional practice: An example from mathematics and science. *Educational Evaluation and Policy Analysis, 25*(1), 1–29.

Hannan, M. T., & Freeman, J. (1984). Structural Inertia and Organizational Change. *American Sociological Review*, *49*(2), 149. http://doi.org/10.2307/2095567

Hannaway, J., & Sproull, L. S. (1978). Who's running the show? Coordination and control in educational organizations. *Administrator's Notebook*, *27*(9), 1-4.

Hargreaves, A., & Goodson, I. (2006). Educational change over time? The sustainability and non-sustainability of three decades of secondary school change and continuity. *Educational Administration Quarterly, 42*(1), 3-41.

Hargreaves, A. & Shaw, P. (2000*). Change Frames: Supporting Secondary Teachers in Interpreting and Integrating Secondary School Reform*. Toronto, Canada: OISE/UT and the Peel-University Partnership.

Harris, A. (2007). The current crisis in leadership: Threat or opportunity. *School Leadership and Management, 27* (2), 105-107.

Harris, A., & Jones, M. (2010). Professional learning communities and system improvement. *Improving Schools*, *13*(2), 172-181.

Hess, F. M., & Fullerton, J. (2009). The numbers we need: Bringing balanced scorecards to education data. *Phi Delta Kappan, 90*(9), 665–69.

Hickey, D. T., Moore, A. L., & Pellegrino, J. W. (2001). The motivational and academic consequences of elementary mathematics environments: Do constructivist innovations and reforms make a difference? *American Educational Research Journal, 38*(3), 611–652.

Hightower, A. M. (2002). *San Diego's big boom: District bureaucracy supports culture of learning (No. R-02–2)*. Seattle, WA: Center for the Study of Teaching and Policy, University of Washington.

Hökkä, P., & Vähäsantanen, K. (2013). Agency-centred coupling—a better way to manage an educational organization? *International Journal of Leadership in Education*, *17*(2), 131–153. http://doi.org/10.1080/13603124.2013.783932

Hollins, E. R., McIntyre, L. R., DeBose, C., Hollins, K. S., & Towner, A. (2004). Promoting a self-sustaining learning community: Investigating an internal model for teacher development. *International Journal of Qualitative Studies in Education*, *17*(2), 247-264.

Honig, M. I. (2003). Building policy from practice: District central office administrators' roles and capacity for implementing collaborative education policy. *Educational Administration Quarterly, 39*(3), 292-338.

Honig, M. I. (2006). Street-level bureaucracy revisited: Frontline district central office administrators as boundary spanners in education policy implementation. *Educational Evaluation and Policy Analysis, 28*(4), 357-383.

Honig, M. I. (2009). No small thing: School district central office bureaucracies and the implementation of New Small Autonomous Schools Initiatives. *American Educational Research Journal, 46*(2), 387-422.

Honig, M. I. (2012). District central office leadership as teaching how central office administrators support principals' development as instructional leaders. *Educational Administration Quarterly, 48*(4), 733-774.

Honig, M. I., & Coburn, C. (2008). Evidence-based decision making in school district central offices: Toward a policy and research agenda. *Educational Policy, 22*(4), 578-608.

Honig, M. I., & Hatch, T. C. (2004). Crafting coherence: How schools strategically manage multiple, external demands. *Educational Researcher*, *33*(8), 16-30.

Hood, C. (1991). A public management for all seasons?. *Public administration*, *69*(1), 3-19.

Hord, S. (1997). *Professional learning communities: Communities of continuous inquiry and improvement*. Austin, TX: Southwest Educational Development Laboratory.

Hox, J.J. & Maas, C. J. M. (2001). The accuracy of multilevel structural equation modeling with pseudo-balanced groups and small samples. *Structural Equation Modeling*, *8*(2), 157-174.

Hubbard, L., Mehan, H., & Stein, M. K. (2006). *Reform as learning: School reform, organizational culture, and community politics in San Diego*. New York, NY: Routledge.

Ikemoto, G.S. & Marsh, J.A. (2007). Cutting through the "data-driven" mantra: Different conceptions of data-driven decision making. *Yearbook of the National Society for the Study of Education, 106*(1), 105-131.

Ingersoll, R. (1993). Loosely coupled organizations revisited. *Research in the Sociology of Organizations, 11*, 81-112.

Ingersoll, R. (1994). Organizational control in secondary schools. *Harvard Review, 64*(2), 150-172.

Ingersoll, R. (1996). Teachers' decision-making power and school conflict. *Sociology of Education*, 159-176.

Ingram, D., Louis, K. S., & Schroeder, R. G. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *Teachers College Record, 106*(6), 1258–1287.

Jacob, B. A. (2005). Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools. *Journal of public Economics*, *89*(5), 761-796.

Jennings, J. (2012a). Reflections on a half-century of school reform: Why have we fallen short and where do we go from here?. Washington, DC: Center on Education Policy.

Jennings, J. (2012b). The effects of accountability system design on teachers' use of test score data. *Teachers College Record, 114*(11), 1-12.

Jimerson, J. B. (2014). Thinking about data: Exploring the development of mental models for "data use" among teachers and school leaders. *Studies in Educational Evaluation*, *42*, 5–14. http://doi.org/10.1016/j.stueduc.2013.10.010

Jimerson, J. B., & McGhee, M. W. (2013). Leading inquiry in schools: Examining mental models of data-informed practice. *Current Issues in Education*, *16*(1). 1-21.

Johnson, L., Smith, R., Willis, H., Levine, A. & Haywood, K. (2011). *The 2011 Horizon Report.* Austin, TX: The New Media Consortium.

Johnson, P. E., & Chrispeels, J. H. (2010). Linking the Central Office and Its Schools for Reform. *Educational Administration Quarterly, 46*(5), 738–775. http://doi.org/10.1177/0013161X10377346

Katz, F. (1964). The school as a complex social organization: A consideration of patterns of autonomy. *Harvard Educational Review*, *34*(3), 428-455.

Kauermann, G., & Carroll, R. J. (2001). A note on the efficiency of sandwich covariance matrix estimation. *Journal of the American Statistical Association*, *96*(456), 1387-1396.

Knapp, M. S. (1997). Between systemic reforms and the mathematics and science classroom: The dynamics of innovation, implementation, and professional learning. *Review of Educational Research*, *67*(2), 227-266.

Kreysing, M. (2002). Autonomy, accountability, and organizational complexity in higher education: the Goettingen model of university reform. *Journal of Educational Administration*, *40*(6), 552-560.

Krieg, J. M. (2008). Are students left behind? The distributional effects of the No Child Left Behind Act. *Education Finance and Policy, 3*(2), 250– 281.

Lachat, M. A., Williams, M., & Smith, S. C. (2006). Making sense of all your data. *Principal Leadership. (Middle Level Ed.). 7*(2), 16-21. Retrieved from http://proquest.umi.com/pqdweb?did=1152250201&Fmt=3&clientId=1585&RQT=309&VName=PQD

Lauen, D. L., & Gaddis, S. M. (2012). Shining a Light or Fumbling in the Dark? The Effects of NCLB's Subgroup-Specific Accountability on Student Achievement. *Educational Evaluation and Policy Analysis*, *34*(2), 185–208. http://doi.org/10.3102/0162373711429989

Lee, J., & Wong, K. K. (2004). The impact of accountability on racial and socioeconomic equity: Considering both school resources and achievement outcomes. *American Educational Research Journal, 41*, 797–832

Leithwood, K., & Levin, B. (2005). Assessing leadership effects on student learning: Selected challenges for research and program evaluation. In W. K. Hoy & C. G. Miskel (Eds.), *Educational leadership and reform* (pp. 53–76). Greenwich, CT: Information Age Publishing.

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

Little, J. W. (2012). Understanding data use practice among teachers: The contribution of micro-process studies. *American Journal of Education, 118* (2), 143-166.

Little, J.W. (1999). *Teachers professional development in the context of high school reform: Findings from a three-year study of restructuring schools*. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, April.

Loeb, S., & Strunk, K. (2007). Accountability and local control: Response to incentives with and without authority over resource generation and allocation. *Education Finance and Policy, 2*(1), 10-39.

Lortie, D. C. (1975). *Schoolteacher: A sociological study*. Chicago, IL: University of Chicago Press.

Louis, K. S. (2007). Changing the culture of schools: Professional community, organizational learning, and trust. *Journal of school leadership*, *16*, 477-487.

Louis, K. S., & Marks, H. M. (1998). Does professional community affect the classroom? Teachers' work and student experiences in restructuring schools. *American Journal of Education*, 532-575.

Louis, K. S., Leithwood, K., Wahlstrom, K. L., & Anderson, S. E. (2010). *Investigating the links to improved student learning: Final report of research findings*. Minneapolis: University

of Minnesota. Retrieved from http://www.cehd.umn.edu/carei/Leadership/Learning-from-Leadership_Final-Research-Report_July-2010.pdf

Luthans, F. (2010). *Organisational Behaviou*r, 12th ed. New York, NY: McGraw-Hill.

Malen, B. (2011). An enduring issue: The relationship between political democracy and educational effectiveness. In D. Mitchell, R. Crowson, & D. Shipps (Eds.), *Shaping education policy: Power and process* (pp. 23-60). New York, NY: Routledge

Mandinach, E. B. (2012). A perfect time for data use: Using data-driven decision making to inform practice. *Educational Psychologist*, *47*(2), 71-85.

Mandinach, E. B., Honey, M., & Light, D. (2006, April). *A theoretical framework for data driven decision making.* Paper presented at the annual meeting of American Educational Research Association, San Francisco. Retrieved from http://cct.edc. org/admin/publications/speeches/DataFrame_AERA06.pdf

March, J. (1978). American public school administrator. A short analysis. *School Review,* 86, 217-250.

March, J. G. (1987). Ambiguity and accounting: The elusive link between information and decision making. *Accounting, Organizations, and Society, 12(2)*, 153-168.

Marsh, J. A., Pane, J. F., & Hamilton, L. S. (2006). *Making sense of data-driven decision making in education: Evidence from recent RAND research (OP-170).* Santa Monica, CA: RAND Corporation.

Mason, S. (2002). *Turning data into knowledge: Lessons from six Milwaukee public schools (WCERWorking Paper No. 2002-3).* Retrieved from: http://www.wcer.wisc.edu/publications/workingpaper/paper/Working_Paper_No_2002_3.pdf

Mason, S. A. (2003). *Learning from data: The role of professional learning communities*. Paper presented at the 2003 Annual Meeting of the American Educational Research Association, Chicago IL.

Massell, D., Kirst, M., & Hoppe, M. (1997). *Persistence and Change: Standards-Based Systemic Reform in Nine States*. CPRE Policy Briefs, RB-21.

Mathis, W. J. (2010). *The "Common Core" standards initiative: An effective reform tool?* Boulder, CO: National Educational Policy Center. Retrieved from http://epicpolicy.org/publication/common-core -standards

McCoach, D. B. (2010). Hierarchical linear modeling. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (pp. 123–140). New York, NY: Routledge.

McCombs, J. S., Kirby, S. N., Barney, H., Darilek, H., & Magee, S. J. (2004). *Achieving state and national literacy goals, a long uphill road: A report to Carnegie Corporation of New York.* Santa Monica, CA: RAND Corporation. Retrieved from http://www.rand.org/pubs/technical_reports/TR180-1/

McLaughlin, M. W., & Mitra, D. (2001). Theory-based change and change-based theory: Going deeper and going broader. *Journal of Educational Change, 2*(4), 301–323.

McLaughlin, M. W., & Talbert, J. E. (2006). *Building school-based teacher learning communities: Professional strategies to improve student achievement.* New York, NY: Teachers College Press

McLaughlin, M.W. (1987) Learning from experience: Lessons from policy implementation. *Educational Evaluation and Policy Analysis 9* (2), 171–178.

McNamara, J. F. (2000). Teaching statistics in principal preparation programs: Part 2. *International Journal of Educational Reform*, *9*(4), 373-384.

Means, B., Padilla, C., & Gallagher, L. (2010) *Use of education data at the local level: From accountability to instructional improvement.* Washington, DC: U.S. Department of Education, Office of Planning, Evaluation, and Policy Development.

Means, B., Padilla, C., DeBarger, A., & Bakia, M. (2009). *Implementing data-informed decision-making in schools: Teacher access, supports, and use.* Washington, DC: U.S. Department of Education, Office of Planning, Evaluation, and Policy Development.

Meyer, H. D. (2002a). The new managerialism in education management: corporatization or organizational learning?. *Journal of Educational Administration*, *40*(6), 534-551.

Meyer, H. D. (2002b). From "loose coupling" to "tight management"? Making sense of the changing landscape in management and organization theory. *Journal of Educational Administration*, *40*(6), 515-520.

Meyer, J. W., & Rowan, B. (1978). The Structure of educational organizations. In Meyer, J. W. (Ed.), *Environments and Organizations* (pp. 78-109). San Francisco, CA: Jossey-Bass.

Morley, L., & Rassool, N. (1999). *School effectiveness: Fracturing the discourse*. London: Falmer Press.

Morley, L., & Rassool, N. (2000). School effectiveness: New managerialism, quality and the Japanization of education. *Journal of Education Policy*, *15*(2), 169–183. http://doi.org/10.1080/026809300285881

Murphy, J., & Datnow, A. (2003). Leadership lessons from comprehensive school reform designs. In J. Murphy & A. Datnow (Eds), *Leadership for school reform: Lessons from comprehensive school reform designs* (pp. 263–278). Thousand Oaks, CA: Corwin Press.

National Research Council. (1997). *Educating one and all: Students with disabilities and standards-based reform*. Washington, DC: National Academy Press.

Neal, D., & Schanzenbach, D. W. (2010). Left behind by design: Proficiency counts and test-based accountability. *The Review of Economics and Statistics*, *92*(2), 263-283.

Nelson, T. H., & Slavit, D. (2007). Collaborative inquiry among science and mathematics teachers in the USA: professional learning experiences through cross-grade, cross-discipline dialogue. *Journal of in-service education, 33*(1), 23-39.

O'Day, J. A. (2002). Complexity, accountability, and school improvement. *Harvard Educational Review*, *72*(3), 293-329.

Orton, J. D., & Weick, K. E. (1990). Loosely coupled systems: A reconceptualization. *Academy of Management Review*, *15*(2), 203-223.

Pajak, E., & Green, A. (2003). Loosely coupled organizations, misrecognition, and social reproduction. *International Journal of Leadership in Education*, *6*(4), 393-413.

Pardini, P. (2000). Data, well done. *Journal of Staff Development*, *21*(1), 12-18.

Phillips, J. (2003). Powerful learning: Creating learning communities in urban school reform. *Journal of curriculum and Supervision*, *18*(3), 240-258.

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

Puma, M., Raphael, J., Olson, K., & Hannaway, J. (2000). *Putting standards to the test: A design for evaluating the systemic reform of education*. Washington, DC: Urban Institute. Retrieved from http://www.urban.org/Uploadedpdf/410385_Standards.pdf

Quellmalz, E., Shields, P. M., & Knapp, M. S. (1995). *School-based reform: Lessons from a national study: A guide for school reform teams.* Washington, DC: U.S. Department of Education, Office of the Under Secretary, Planning and Evaluation Service

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. 2nd edition. Newbury Park, CA: Sage.

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

Reback, R., Rockoff, J., & Schwartz, H. L. (2011). *Under pressure: Job security, resource allocation, and productivity in schools under NCLB* (Working paper). The National Bureau of Economic Research, NBER Working Paper No. 16745.

Reeves, P. L., & Burt, W. L. (2006). Challenges in data-based decision-making: Voices from principals. *Educational Horizons*, *85*(1), 65-71.

Roderick, M. (2012). Drowning in data but thirsty for analysis. *Teachers College Record*, *114*(11), 1-9.

Rosenholtz, S. (1989). *Teacher's workplace: The social organization of schools*. New York: Longman.

Rosenshine, B. (2003). High-stakes Testing: Another Analysis. *Education Policy Analysis Archives, 11*(24), 11-10. http://doi.org/10.14507/epaa.v11n24.2003

Rosenthal, R., & Rosnow, R. L. (1984). *Essentials of behavioral research: Methods and data analysis*. New York: McGraw-Hill.

Rothstein, R. (2008). Leaving “No Child Left Behind” behind. *American Prospect, 19*(1), 50–54.

Rowan, B. (2002). Rationality and reality in organizational management: Using the coupling metaphor to understand educational (and other) organizations-a concluding comment. *Journal of Educational Administration*, *40*(6), 604-611.

Rudner, L. M., & Boston, C. (2003). Data warehousing: Beyond disaggregation. *Educational Leadership, 60*(5), 62-65.

Russell, J. L., Meredith, J., Childs, J., Stein, M. K., & Prine, D. W. (2015). Designing Inter-Organizational Networks to Implement Education Reform: An Analysis of State Race to the Top Applications. *Educational Evaluation and Policy Analysis*, *37*(1), 92–112. http://doi.org/10.3102/0162373714527341

Sanchez, R., & Mahoney, J. T. (1996). Modularity, flexibility, and knowledge management in product and organization design. *Strategic management journal*, *17*(S2), 63-76.

Sarason, S. (1982). *The culture of the school and the problem of change* (2nd ed.). Boston, MA: Allyn & Bacon.

Sarason, S. (1990). *The predictable failure of educational reform*. San Francisco, CA: Jossey-Bass.

Schmoker, M., & Marzano, R. J. (1999). Realizing the promise of standards-based education. *Educational Leadership*, *56*(6), 17–21.

Senge, P. M. (1990). The fifth discipline: The art and practice of the learning organization. New York, NY: Doubleday/Currency.

Sheldon, K., & Biddle, B. (1998). Standards, accountability, and school reform: Perils and pitfalls. *The Teachers College Record.100*(1), 164-180.

Shen, J., & Cooley, V. E. (2008). Critical issues in using data for decision-making. *International Journal of Leadership in Education*, *11*(3), 319–329. http://doi.org/10.1080/13603120701721839

Shen, J., & Ma, X. (2006). Does systemic change work? Curricular and instructional practice in the context of systemic change. *Leadership and Policy in Schools*, *5*(3), 231-256.

Shen, J., Cooley, V., Reeves, P., Burt, W., Ryan, L., Rainey, J. M., & Yuan, W. (2010). Using data for decision-making: Perspectives from 16 principals in Michigan, USA. *International Review of Education, 56*(4), 435-456.

Shen, J., Palmer, L., & Crawford, C. (2005). The importance of educational goals as perceived by principals. *School Principals*, 46-60.

Smith, M. S., & O'Day, J. A. (1990). Systemic School Reform. In S. Fuhrman & B. Malen (Eds.), *The politics of curriculum and testing* (pp. 233-267). Philadelphia, PA: The Falmer Press.

Spillane, J. P. (1994). How districts mediate between state policy and teachers' practice. In R. F. Elmore & S. H. Fuhrman (Eds.), *The governance of curriculum: 1994 yearbook of the Association for Supervision and Curriculum Development* (pp. 167–185). Alexandria, VA: Association for Supervision and Curriculum Development

Spillane, J. P. (1998). State Policy and the Non-Monolithic Nature of the Local School District: Organizational and Professional Considerations. *American Educational Research Journal*, *35*(1), 33–63. http://doi.org/10.3102/00028312035001033

Spillane, J. P. (1999). External reform initiatives and teachers' efforts to reconstruct their practice: The mediating role of teachers' zones of enactment. *Journal of Curriculum Studies*, *31*(2), 143–175.

Spillane, J. P. (2000). Cognition and policy implementation: District policymakers and the reform of mathematics education. *Cognition and Instruction, 18*(2), 141–179.

Spillane, J. P. (2012). Data in practice: conceptualizing the data-based decision-making phenomena. *American Journal of Education*, *118*(2), 113–141.

Spillane, J. P., & Burch, P. (2003). *Policy, administration, and instructional practice: Loose coupling revisited.* Retrieved from www.northwestern.edu/ipr/publications/papers/2003/WP-03–04.pdf

Spillane, J. P., & Jennings, N. E. (1997). Aligned instructional policy and ambitious pedagogy: Exploring instructional reform from the classroom perspective. *Teachers College Record, 98*(3), 449–481.

Spillane, J. P., Parise, L. M., & Sherer, J. Z. (2011). Organizational routines as coupling mechanisms policy, school administration, and the technical core. *American Educational Research Journal*, *48*(3), 586-619.

Stecher, B. & Barron, S. (1999). *Test-Based Accountability: The perverse consequences of milepost testing.* Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Canada, April 1999

Stevenson, D. L., & Schiller, K. S. (1999). State education policies and changing school practices: Evidence from the national longitudinal study of schools, 1980-1993. *American Journal of Education*, *107*(4), 261-288. http://doi.org/10.2307/1085710

Stoll, L., Bolam, R., McMahon, A., Wallace, M., & Thomas, S. (2006). Professional learning communities: A review of the literature. *Journal of educational change*, *7*(4), 221-258.

Strahan, D. (2003). Promoting a collaborative professional culture in three elementary schools that have beaten the odds. *The Elementary School Journal*, 127-146.

Supovitz, J. (2009). Can high stakes testing leverage educational improvement? Prospects from the last decade of testing and accountability reform. *Journal of Educational Change*, *10*(2-3), 211-227.

Supovitz, J. A., & Klein, V. (2003). *Mapping a course for improved student learning: How innovative schools systematically use student performance data to guide improvement*. Philadelphia, PA: Consortium for Policy Research in Education.

Sutherland, S. (2004). Creating a culture of data use for continuous improvement: A case study of an Edison Project school. *American Journal of Evaluation*, *25*(3), 277-293.

Swanson, C. B., & Stevenson, D. L. (2002). Standards-based reform in practice: Evidence on state policy and classroom instruction from the NAEP state assessments. *Educational Evaluation & Policy Analysis, 24*(1), 1–27.

Thorn, C. A. (2001). Knowledge Management for Educational Information Systems: What Is the State of the Field?. *Education policy analysis archives,9*(47), Retrieved from http://epaa.asu.edu/epaagv9n47/.

Thornton, B., & Perreault, G. (2002). Becoming a data-based leader: An introduction. *NASSP Bulletin, 86*(630), 86-96.

Togneri, W. & Anderson, S. E. (2003). *Beyond islands of excellence: What districts can do to improve instruction and achievement in all schools*. Washington, D.C.: The Learning First Alliance and the Association for Supervision and Curriculum Development.

U.S. Department of Education (2009), *Race to the Top program: Executive summary*. U.S. Department of Education, Washington, DC, Retrived from www2.ed.gov/programs/racetothetop/executive-summary.pdf

U.S. Department of Education. (2006). *No Child Left Behind is working*. Retrieved from http://www.ed.gov/nclb/overview/importance/nclbworking.html.

Wang, M. & Long, Q. (2011). Modified robust variance estimator for generalized estimating equations with improved small-sample performance. *Statistics in Medicine, 30*(11),1278-1291.

Webb, N. L. (2005). *Alignment analysis of mathematics standards and assessments, Michigan High School*. Madison, WI: University of Wisconsin.

Webb, N. L. (2006). *Alignment analysis of mathematics standards and assessments, Tennessee Grades 3–9*. Madison, WI: University of Wisconsin.

Wei, X. (2012a). Are more stringent NCLB State accountability systems associated with better student outcomes? An analysis of NAEP results across states. *Educational Policy, 26*(2), 268-308.

Wei, X. (2012b). Does NCLB improve the achievement of students with disabilities? A regression discontinuity design. *Journal of Research on Educational Effectiveness*, *5*(1), 18-42.

Weick, K. E. (1976). Educational organizations as loosely-coupled systems. *Administrative Science Quarterly, 21*(l): 1-19.

Willms, J. D. (2000). Monitoring school performance for'standards-based reform'. *Evaluation & Research in Education, 14*(3-4), 237-253.

Wong, M., Cook, T. D., & Steiner, P. M. (2009). *No Child Left Behind: An interim evaluation of its effects on learning using two interrupted time series each with its own non-equivalent comparison series (Working paper).* Chicago, IL: Northwestern University Institute for Policy Research.

Xia, J., Gao, X., & Shen, J. (2015). Demographic and Professional Characteristics of Urban School Principals in the United States. In M. Khalifa, N. W. Arnold, A. F. Osanloo, & C. M. Grant (Eds.). *Handbook of Urban Educational Leadership* (pp. 23-42). Lanham, MD: Rowman & Littlefield.

Young, V. M. (2006). Teachers' use of data: Loose coupling, agenda setting, and team norms. *American Journal of Education, 112*(4), 521-548.

Appendix A

Descriptive Statistics

Variable	Representation (Name in the SASS data)	Coding	Statistics
Teacher-level Outcome			
TDIDM	Teacher Data Informed Decision Making, Cronbach's alpha = .992, Recoded as Mean of T0287, T0288, T0289	Continuous	M = 2.17, SD = 1.27
Teacher background (as teacher level variables)			
Level	Teaching level (TLEV2_03) by NCES	Elementary = 1	16.10%
		Secondary = 2	83.90%
Experience	Total teaching experience (TOTEXPER) by NCES	Continuous	M = 14.18, SD = 10.23
Gender	Gender (T0408)	Male = 1	24.50%
		Female = 0	75.50%
Field	Teaching field (ASSIGN03)	Other than Maths/science = 0	85.70%
		Maths/science = 1	14.30%
Cert	Certification (T0166)	Provisional/temporary/emergency/none = 0	8.50%
		Regular/probationary = 1	91.50%
School Background (as school level variables)			
Stuprogres	Reports used to evaluate stud progress (S0457),	Valid skip & no = 0	4.40%
		yes = 1	95.60%
YrFocus	Reports used for next year focus (S0458)	Valid skip & no = 0	8.10%
		yes = 1	91.90%
RealignCurrlm	Reports used to realign curriculum (S0459)	Valid skip & no = 0	10%
		yes = 1	90%

InformParnt	Reports used to inform parent (S0460)	Valid skip & no = 0	4%
		yes = 1	96%
PrmptImpve	Reports used prompt improvement (S0461)	Valid skip & no = 0	6.60%
		yes = 1	93.40%
SDIDM	Recoded as Sum of Stuprogres, YrFocus, RealignCurrlm, InformParnt, PrmptImpve, Cronbach's alpha = .825	no report =0	2.20%
		1 area =1	0.60%
		2 areas = 2	1.30%
		3 areas = 3	2.90%
		4 areas = 4	8.40%
		5 areas = 5	80.10%
TSexp	Aggregated School Mean of TOTEXPER by NCES	Continuous	M = 14.64, SD = 6.06
Tscert	Aggregated School Mean of Certification (T0166)	Continuous	M = .92, SD = 0.17
Tsgender	Aggregated School Mean of Teacher Gender (T0408)	Continuous	M = .29, SD = .28
HighTeach	Teachers teaching to high academic standards (A0149)	Continuous	M = 81.34, SD = 18.04
GroupStudent	Aggregated School Mean of T0287	Continuous	M = 1.99, SD = 0.74
AssessAreas	Aggregated School Mean of T0288	Continuous	M = 2.77, SD = 0.72
AdjustCurriculum	Aggregated School Mean of T0289	Continuous	M = 2.86, SD = 0.69
School-level Outcome			
SchPass	School performance (A0166)	Pass none/some/most = 0	45.70%
		Pass all = 1	54.30%

Principal background (as school level variables)

Experience	Principal experience (A0025)	Continuous	M = 7.76, SD = 7.09
Enrol100	Total K-12 and ungraded enrollment divided by100 (ENRK12UG), created by NCES from school data	Continuous	M = 5.40, SD = 4.37
MINOR	Percentage of students in school who were of a racial/ethnic minority (MINENR) by NCES	Continuous	M = 38.76, SD = 35.30
NSLAP	Percentage of enrolled students approved for the National School Lunch Program at school (NSLAPP_S) by NCES	Continuous	M = 44.65, SD = 28.05
Locat1	Urbanicity of the school (URBANS03) by NCES	Urban fringe of a large or mid-size central city & Small town/rural = 0	75.10%
		Large or mid-size central city = 1	24.90%
Locat2	Urbanicity of the school (URBANS03) by NCES	Large or mid-size central city & Small town/rural= 0	51.90%
		Urban fringe of a large or mid-size central city= 1	48.10%
Schperform	School performance (A0166)	No standard/Pass all = 0	60.40%
		Pass most/some/no = 1	39.60%
Gender	Gender (A0254)	Female = 0	68.20%
		Male = 1	31.80%
SCHLevl1	School level (SCHLEVEL) by NCES	Secondary & combined = 0	29.80%
		Elementary = 1	70.20%
SCHLevl2	School level (SCHLEVEL) by NCES	Elementary & combined = 0	77.50%
		Secondary = 1	22.50%

DIDMTEST	Assess-state or national tests (A0161)	Valid skip & no = 0	13.00%
		yes = 1	87.00%
DIDMportfolio	Assess-student portfolios (A0162)	Valid skip & no = 0	51.70%
		yes = 1	48.30%
DIDMSurvey	Assess-parent or student surveys (A0163)	Valid skip & no = 0	24.70%
		yes = 1	75.30%
PDS	Recoded as Sum of data-informed decision-makingTEST, data-informed decision-makingportfolio, data-informed decision-makingSurvey, Cronbach's alpha = .615	no report =0	2.20%
		none = 0	10.90%
		1 stream = 1	9.00%
		2 streams = 2	38.60%
		2 streams = 3	41.40%
District Background (as District level variables)			
DStuprogres	Reports used to evaluate stud progress (D0147),	Valid skip & no = 0	6.00%
		yes = 1	94.00%
DYrFocus	Reports used for next year focus (D0458)	Valid skip & no = 0	8.80%
		yes = 1	91.20%
DRealignCurrlm	Reports used to realign curriculum (D0459)	Valid skip & no = 0	7%
		yes = 1	93%
DInformParnt	Reports used to inform parent (D0460)	Valid skip & no = 0	5%
		yes = 1	95%
DPrmptImpve		Valid skip & no = 0	7.40%

	Reports used prompt improvement (D0461)	yes = 1	92.60%
		no report =0	3.90%
DDIDM	Recoded as Sum of DStuprogres, DYrFocus, DRealignCurrlm, DInformParnt, DPrmptImpve, Cronbach's alpha = .913	1 area =1	0.60%
		2 areas = 2	0.90%
		3 areas = 3	1.50%
		4 areas = 4	6.00%
		5 areas = 5	87.00%

Appendix B

SASS Questions for TDIDM, PDS, SDIDM, and DDIDM

SASS Questions Measuring the Extent of Teacher's Data Informed Decision Making

54. **Do you have access to your students' scores on state or district achievement tests?**

0286

1 ☐ Yes

2 ☐ No

3 ☐ Not applicable; my students haven't taken a state or district achievement test.

GO TO item 56 below.

55. **To what extent do you use the information from your students' state or district achievement test scores –**

		Mark (X) ONE box on each line.			
		Not at all	To a small extent	To a moderate extent	To a great extent
a. To group students into different instructional groups by achievement or ability?	0287	1 ☐	2 ☐	3 ☐	4 ☐
b. To assess areas where you need to strengthen your content knowledge or teaching practice?	0288	1 ☐	2 ☐	3 ☐	4 ☐
c. To adjust your curriculum in areas where your students encountered problems?	0289	1 ☐	2 ☐	3 ☐	4 ☐

SASS Questions Measuring Principal's Sources of Data for Assessing Schools

26a. **Does this school have a formal school improvement plan?**

0160

1 ☐ Yes

2 ☐ No → *GO TO item 27a below.*

b. **Do you use any of the following to assess this school's progress on that plan?**

(1) **State or national tests**

0161

1 ☐ Yes

2 ☐ No

(2) **Parent or student surveys**

0162

1 ☐ Yes

2 ☐ No

(3) **Student portfolios**

0163

1 ☐ Yes

2 ☐ No

SASS Questions Measuring the Extent of School's Data Informed Decision Making

19. Does this school receive performance reports from the district that cover such things as students' scores on achievement tests or graduation rates?

0455 1 ☐ Yes → *GO TO item 21 on page 9.*

2 ☐ No

21. Does this school use these performance reports to –

a. Evaluate the progress of students in this school?

0457 1 ☐ Yes

2 ☐ No

b. Determine the next year's instructional focus?

0458 1 ☐ Yes

2 ☐ No

c. Realign the curriculum, such as with content standards and/or other indicator criteria?

0459 1 ☐ Yes

2 ☐ No

d. Inform parents and the community of the school's progress?

0460 1 ☐ Yes

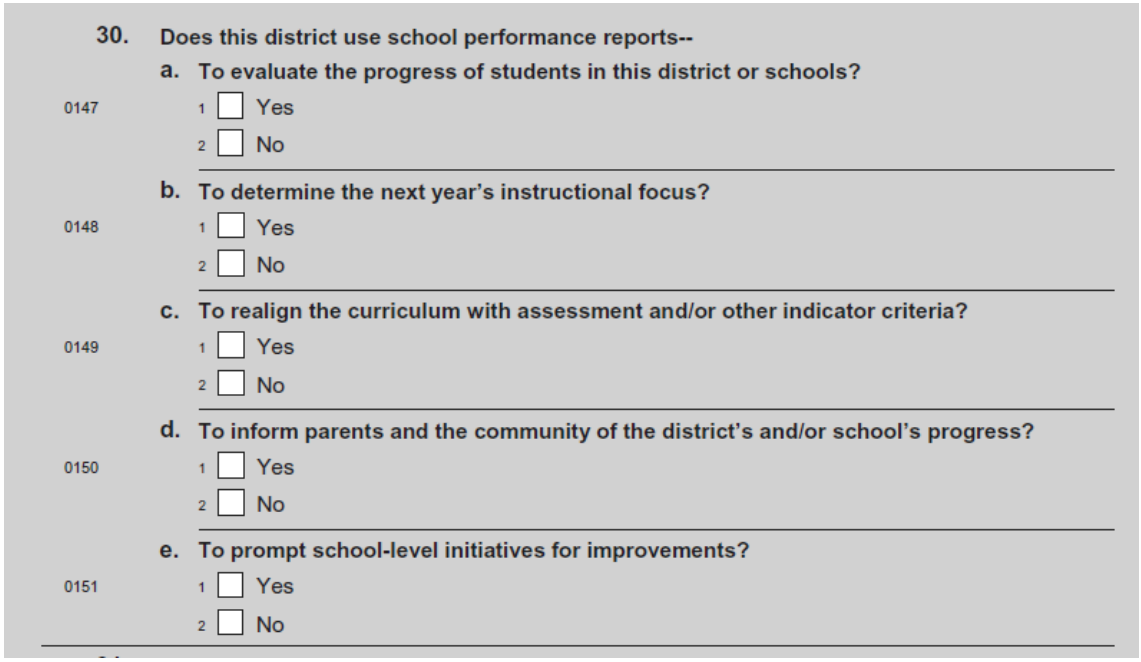
2 ☐ No

e. Prompt school-level initiatives for improvement?

0461 1 ☐ Yes

2 ☐ No

SASS Questions Measuring the Extent of District's Data Informed Decision Making



30. Does this district use school performance reports--

a. To evaluate the progress of students in this district or schools?

0147 1 ☐ Yes
2 ☐ No

b. To determine the next year's instructional focus?

0148 1 ☐ Yes
2 ☐ No

c. To realign the curriculum with assessment and/or other indicator criteria?

0149 1 ☐ Yes
2 ☐ No

d. To inform parents and the community of the district's and/or school's progress?

0150 1 ☐ Yes
2 ☐ No

e. To prompt school-level initiatives for improvements?

0151 1 ☐ Yes
2 ☐ No

Appendix C

HSIRB Approval Letter

WESTERN MICHIGAN UNIVERSITY

Human Subjects Institutional Review Board

Date: April 16, 2014

To: Jianping Shen, Principal Investigator
Jiangang Xia, Student Investigator
Xingyuan Gao, Student Investigator

From: Amy Naugle, Ph.D., Chair Amy Naugle

Re: HSIRB Project Number 14-04-17

This letter will serve as confirmation that your research project titled "Teachers' and Principals' Leadership and Decision-Making Power" has been **approved** under the **exempt** category of review by the Human Subjects Institutional Review Board. The conditions and duration of this approval are specified in the Policies of Western Michigan University. You may now begin to implement the research as described in the application.

Please note: This research may **only** be conducted exactly in the form it was approved. You must seek specific board approval for any changes in this project (e.g., ***you must request a post approval change to enroll subjects beyond the number stated in your application under "Number of subjects you want to complete the study***)." Failure to obtain approval for changes will result in a protocol deviation. In addition, if there are any unanticipated adverse reactions or unanticipated events associated with the conduct of this research, you should immediately suspend the project and contact the Chair of the HSIRB for consultation.

Reapproval of the project is required if it extends beyond the termination date stated below.

The Board wishes you success in the pursuit of your research goals.

Approval Termination: April 15, 2015

251 W. Walwood Hall, Kalamazoo, MI 49008-5456
PHONE: (269) 387-8293 FAX: (269) 387-8276