

TEACHERS “TALK BACK”: EXPLORING THE DYNAMICS BETWEEN
PRACTICE AND VALUE-ADDED EVALUATION POLICY

by

Emily Astor

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Doctor of Education Degree
in Assessment, Learning and School Improvement

Middle Tennessee State University

August 2018

Dissertation Committee:

Dr. Rick Vanosdall, Chair

Dr. Heather Dillard

Dr. Nina Gregg

Dr. Kevin Krahenbuhl

ACKNOWLEDGMENTS

I would like to thank my committee members Drs. Heather Dillard and Kevin Krahenbuhl, who provided invaluable guidance and encouragement throughout the writing process. As a member of the first Assessment, Learning, and School Improvement cohort, I was fortunate enough to write this dissertation with the support of nineteen exceptional educators, administrators, and researchers. The friendships and collegial relationships that developed over the course of my doctoral program at Middle Tennessee State University challenged and supported me during a time of transformation and professional growth.

I am grateful for the unending grace with which my sons, Henry and Teddy, endured many weekends, holidays, and evenings without my full presence while I was attending class, completing coursework, and writing my dissertation. Without fail, my family was flexible and patient with my schedule. They gifted me with the time I needed to travel to campus three hours away and to meet deadlines. I am especially grateful to Aaron and Samantha Astor. I could not ask for better partners in parenting.

I am forever indebted to my mother and father, Mary and Siegfried Hausner. Their confidence in my abilities empowered me to begin this degree and their generosity allowed me to complete it. I hope I have shown my gratitude by passing along to my sons the same love, affirmation, and support my parents have shown me.

I would like to thank Dr. Nina Gregg for her counsel since the outset of this study. She generously gave her time and attention from the conceptualization of the research questions to the final draft of the dissertation. Through our many editing sessions, my writing has improved, and I have improved as a writer.

My advisor, Dr. Rick Vanosdall, is a model and champion of transformative learning. The course of my professional life has changed under his guidance and I am deeply grateful for

his investment in my work and my learning journey. His vision as a university administrator, program director, professor, and mentor has touched so many lives and I am fortunate to be one of them.

Lastly, I would like to thank the teachers who shared their stories with me. Their commitment to their students' learning and welfare at great personal cost continues to humble me. Their dedication also inspires me to keep listening and sharing their stories with others. So much of our profession is misunderstood, even by those within it. My hope is that the insights of these teachers will be heard and that their wisdom will lead to a deeper understanding of the lives of those we ask so much from - elementary educators.

ABSTRACT

The accountability movement in education since the No Child Left Behind Act of 2001 has produced value-added evaluation policies in the United States that have resulted in discord and undesirable responses among many teachers. Despite investigations into the validity of value-added evaluation policies and descriptive reports of teachers' responses to value-added evaluation, education research has shed little light on the effects of value-added teacher evaluation policy implementation. The study draws from inhabited institutionalism and sensemaking research that suggests teachers and other actors within schools incorporate institutional values and norms into their own practices based on their understanding of the goals or tasks dictated by the policy, their organizational contexts and their professional identity. This study addresses the central question: What theoretical model can explain how teachers have adapted their practices in response to teacher value-added evaluation policy implementation? This qualitative study investigates how teachers adapt their practices in response to value-added teacher evaluation policy by discovering and documenting the self-reported beliefs, perceptions, experiences, and practices related to value-added teacher evaluation among a theoretical sample of teachers (n=19) across a single public-school district in Tennessee. In-depth interviews, member checks, and a process framework was used for analysis. A grounded theoretical model was constructed to describe and relate how (a) the macro-level, structural conditions of teacher evaluation, (b) the phenomena that arose from the structural conditions, and (c) the particular contextual conditions and components of teachers' personal histories interact to influence teachers' adaptations. The study concludes with a discussion of the implications for practice and policy, specifically how teachers, schools and school districts exercise agency in how policy is implemented and how value-added evaluation policy creates a tension between cultural ideologies of accountability and teachers' lived experiences of vulnerability.

Keywords: accountability, value-added, policy implementation, sensemaking, inhabited institutionalism, teacher evaluation

TABLE OF CONTENTS

LIST OF TABLES	viii
LIST OF FIGURES	x
LIST OF APPENDICES	xi
CHAPTER I: INTRODUCTION	1
Problem Statement	5
Purpose of the Study	6
Context	7
Research Questions.....	8
Theoretical and Conceptual Frameworks	8
Summary of Methods and Analysis.....	10
Biases and Assumptions	11
Definition of Terms.....	13
Organization of the Dissertation	17
CHAPTER II: REVIEW OF LITERATURE	18
Review of Study.....	18
Preview of Chapter	18
Conceptual Framework and Theoretical Foundation.....	19
Literature Review of Policy	41
Empirical Research on Value-Added Methods.....	59
CHAPTER III: METHODOLOGY	74
Impetus for Design	74
Rationale for Design	77
Description of Methods.....	89

CHAPTER IV: RESULTS AND ANALYSIS	101
Context of the Study	101
Purpose of the Study	102
Data Collection	103
Demographics of Participants	104
Methodology of Data Analysis	109
Findings and Analysis	111
Analysis of the Four Core Themes	138
Development of Theoretical Model	154
Summary	166
CHAPTER V: CONCLUSIONS	169
Review of Methods	169
Discussion of Findings	171
Implications and Recommendations	199
Suggestions for Future Research	205
Chapter Summary	207
REFERENCES	213
APPENDICES	243
Appendix A:	244
Appendix B:	246

LIST OF TABLES

Table 1. Kelchtermans’ components and descriptions of professional self-understanding	30
Table 2. Race to the Top proposal selection criteria	51
Table 3. Example of calculation of teacher effectiveness score in Tennessee	53
Table 4.1. Participants’ level of elementary education teaching experience	105
Table 4.2. Participants’ past and current employment history by school within the Community School District	106
Table 4.3. Participants’ gender	105
Table 4.4. Participants’ evaluation status	108
Table 4.5. Years of participants’ TVAAS evaluation	108
Table 5.1. Participants’ responses regarding beliefs of the best use of evaluations	114
Table 5.2. Participants’ responses regarding policy intent	116
Table 5.3. Participants’ responses regarding adaptations in practice	120
Table 5.4. Participants’ responses regarding interactions with colleagues	123
Table 5.5. Participants’ responses regarding interactions with principals	128
Table 5.6. Participants’ responses regarding school-wide practices	130
Table 5.7. Participants’ responses regarding issues expressed	134
Table 6.1. Four core themes	139, 154, 171
Table 6.2. Core Theme One with categories	140
Table 6.3. Core Theme Two with categories	142
Table 6.4. Core Theme Three with categories	144
Table 6.5. Core Theme Four with categories	148

Table 7. Teachers' strategies and intended results for improving TVAAS scores193

LIST OF FIGURES

Figure 1. Socio-ecological model adapted from McLeroy et al. (1988)	22
Figure 2. Conceptual framework integrating institutional, sensemaking, and agency theories	35
Figure 3. Detail of conceptual framework integrating institutional, sensemaking, and agency theories.....	36
Figure 4. Theoretical model to explain how teachers choose practices in response to value-added evaluation	157, 172
Figure 5. Teachers' strategies for choosing adaptive practices in response to value-added evaluations	162, 173

LIST OF APPENDICES

Appendix A: Table of Codes

Appendix B: Teacher interview guide for grounded study, Teachers "talk back":

Exploring the dynamics between practice and value-added evaluation.

CHAPTER I

INTRODUCTION

As a result of two independent yet intersecting changes in the policy environment, public school teachers in the United States have become the focus of reform policies designed to improve educational outcomes for all students. Since the 1980s, the narrative about *why* public schools have failed to close achievement gaps shifted from socioeconomic inequality to ineffective teachers. Meanwhile, the prescription for *what* was needed to improve student learning swung from allocating public funds to bring equity across schools to increasing the monitoring of the teacher workforce to ensure its effectiveness. These transitions gained momentum in the public sphere when the United States National Commission on Excellence in Education released its report, *A Nation at Risk* in 1983. Lawmakers from both ends of the political spectrum and the broader policy community began to challenge the ability of federally funded programs to address persistent educational inequality. Advocates of raising student test scores by improving teacher quality argued that ineffective teachers and the tenure policies protecting them were the cause of public education's ills. When Congress reauthorized the Elementary and Secondary Education Act in 2001 with the passage of the No Child Left Behind Act, a "new" accountability paradigm for improving teacher and school effectiveness emerged (Cochran-Smith et al., 2018). Accountability now referred to student and teacher performance, regarded schools and teachers as units of improvement, included public reporting, and attached positive and negative consequences to various performance levels (Fuhrman, 1999). The movement for improving teacher effectiveness gained considerable

attention when President Obama's Race to the Top program of 2009 offered the promise of funding for states that incorporated standardized measures to improve their teacher workforce by evaluating teacher quality and attaching high stakes consequence for ineffective teachers. Unfortunately, the escalating interest in teacher effectiveness and the rise of accountability in education converged to produce education policies that have resulted in controversy and conflict among teachers (Collins, 2014) and stymied progress toward reformers' goal of equity in schools (Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2011).

To improve policy outcomes, a critical review of recent changes in the social and political landscape is necessary to reveal the underlying principles and ideologies of education policy and to understand how these principles and ideologies shape policy implementation. Honig (2006) argued that policy research must identify more than which policies work; research must also seek to uncover the actors and contexts that influence implementation and "provide robust, grounded explanations for how interactions among them help to explain implementation outcomes" (p. 2). By "confronting the complexity of policy implementation" (Honig, 2006, p. 3), research can inform decision-makers with a deeper knowledge of how teachers make sense of evaluation policies and how sensemaking shapes teaching practices in schools and classrooms.

Since World War II, the influence of free market ideologies and scientific management principles have expanded into the administration and delivery of public services of many Western democracies, including public education in the United States. As a result, when U.S. President Johnson's War on Poverty appeared unwinnable through

federal spending programs, the civic optimism of the 1960s evolved into anger and frustration with the impotence of the public sector in the early 1980s. Vexed with the failure of schools to overcome the effects of poverty, reformers shifted their priorities from reducing the negative effects of socioeconomic inequalities on student outcomes to holding schools accountable for student learning. Cochran-Smith et al. (2018) suggested that accountability is now “a major strategy” for improving K-12 education through a discourse of “heightened auditing, monitoring and surveillance” (p. 9). The aim of educational accountability systems has become less concerned with schools’ compliance with policy and more focused on monitoring students’ learning outcomes at the school level (Hallinger, Heck, & Murphy, 2014). During the same time period, the policy environment changed as policymakers took notice of research that singles out teacher effectiveness as the most influential school-level factor in student achievement (Hanushek, 2014; Hattie, 2012; McCaffrey, Lockwood, Koretz, Louis, & Hamilton, 2004b; Rivkin, Hanushek, & Kain, 2005; Stronge, 2007). Educational economist Hanushek (2014) claimed that “teachers have an enormous influence on students and on their futures” (p. 26). Similarly, Rivkin et al. (2005) suggested that high quality instruction in the early years of school can outweigh the negative educational effects associated with low socioeconomic status. Bolstered with compelling quantitative data that suggested that schools can improve student outcomes with better teaching, policymakers designed reforms based upon the belief that teachers, if properly motivated, can overcome the deleterious effects of social and economic inequality on educational outcomes that years of federal funding could not.

As a result of these two changes, many states across the country modified their educator evaluation policies to include systems to score and rank teachers based on their students' standardized test scores, to reward highly effective teachers, and to remove poor ones. In 2017, 39 states required student growth data in teacher evaluations (National Council on Teacher Quality, n.d.) The stated purpose of these evaluation systems is to achieve substantial gains in student learning and to close achievement gaps by recruiting, developing, rewarding and retaining effective teachers (U.S. Department of Education, 2009). Value-added evaluation policies identify teachers as the primary means to improve student outcomes. If teachers are the key to improving student learning, then what they do in their classrooms is instrumental for achieving policy goals.

Education agencies at both the state and federal levels designed evaluation reforms intending to improve student achievement and equity in schools, but their limited understanding of policy implementation has led to value-added evaluation policies that have evoked negative and destructive practices among teachers. Studies by Booher-Jennings (2005) Au (2011), Collins (2014), Hursh (2007), Johnson (2015), and Nichols and Berliner (2007) indicated that value-added evaluation and high stakes standardized testing have resulted in undesirable adaptations in teachers' practices that exacerbate the educational inequalities that the policies were designed to remedy (Booher-Jennings, 2005; Darling-Hammond, 1991; Nichols & Berliner, 2007).

The strong response to value-added evaluation policies has prompted a body of research that includes a range of studies; some defend, and others critique the validity of value-added methods. Sanders and Horn (1994) and Rivkin et al. (2005) claimed that

value-added methods successfully control for students' characteristics and socioeconomic backgrounds to assess teacher influence on student learning. In addition, several institutions including the Statistical Analysis System (SAS) (2015), the RAND Corporation (2002) and the Education Trust (2009) suggested that value-added methods provide valuable data to assess school and teacher effectiveness. Conversely, Darling-Hammond et al. (2011), McCaffrey et al. (2004b), and Piro, Wiemers, and Shutt (2011) claimed that value-added methods cannot control for all biases and warned that teachers' value-added ratings are highly dependent on the students they teach. Other studies have focused on issues with errors in value-added methods. Schochet and Chiang (2013), Corcoran, Jennings, and Beveridge (2011), and the National Research Council (2009) warned that value-added methods produce data that are too unstable and unreliable for teacher evaluation purposes.

Problem Statement

The accountability movement in education today has produced value-added evaluation policies (VAEP) that have provoked discord and undesirable responses among teachers. Despite investigations into the validity of VAEP and teachers' responses to value-added evaluation, education research has shed little light on VAEP implementation thus far. While the body of research critiquing the validity of value-added methods and their applications may help to explain teachers' cynicism toward VAEP, it does not explain *how* or *why* teachers choose their practices when responding to these policies. Similarly, descriptive studies by Collins (2014), Jiang, Sporte, and Luppescu (2015), and

Lee (2011) documented teachers' wide-spread distrust of the validity and purpose of VAEP, but they do not offer explanations of how teachers make their choices in practice when adapting to VAEP. Identifying teachers' perceptions of policy is a start, but it is insufficient: listening to teachers in ways that reveal their understandings of policy within their social contexts, researchers can construct locally-situated accounts to illuminate the complex dynamics of VAEP implementation and ultimately improve education policy design.

Purpose of the Study

The purpose of this empirical study is to explain the process of how teachers respond to value-added evaluation policies (VAEP) in order to understand policy dynamics and teachers' roles in the implementation process. By gathering teachers' narratives and accounts of their beliefs, perceptions, experiences, and practices related to VAEP, this study contributes the voices of teachers navigating the realities of VAEP as empirical evidence of the implications of this policy on teacher behavior. The findings also add grounded research to the theoretical literature on how sensemaking shapes policy implementation. The research is intended to empower teachers, to advise school and district-level administrators who implement evaluation policy, and to enlighten decision-makers who select or design teacher evaluation policy.

Context

In February 2009, U.S. President Barak Obama signed the American Recovery and Reinvestment Act (ARRA) into law. Among other economic stimulation measures, the legislation created a \$4.35 billion Race to the Top (RTTT) fund to reward states that implemented reforms that mandated student growth measurements to evaluate teachers and principals. The selection criteria for RTTT funding stipulated that states included several components in their proposals related to student growth and teacher evaluation. First, states would “establish clear approaches to measuring student growth” (U.S. Department of Education, 2009). Second, states would implement evaluation systems for teachers and principals that “differentiate effectiveness using multiple rating categories that take into account student growth” (U.S. Department of Education, 2009). Third, states would “use these evaluations, at a minimum, to inform decisions regarding ... compensating, promoting, and retaining teachers and principals,” and “whether to grant tenure and/or full certification” (U.S. Department of Education, 2009).

In a special legislative session in January 2010, the Tennessee General Assembly passed the First to the Top Act, a bill requiring school districts to implement Tennessee’s Value-Added Assessment System (TVAAS) and other student outcome measures as a significant component of principal and teacher evaluations (Piro et al., 2011). The First to the Top Act made Tennessee one of the first states in the nation to implement a statewide student outcomes-based, educator evaluation system and helped the state win over \$500 million in RTTT funding (Tennessee Department of Education, 2012). Public school districts across the state implemented a new teacher evaluation system, the Tennessee

Educator Acceleration Model (TEAM), which included TVAAS. While TVAAS had been used in the past to calculate student growth, the 2011-12 school year marked the first time value-added scores were included in teacher evaluations. With the TEAM assessment system, teachers and principals received a TVAAS rating based on their students' annual growth in achievement on the Tennessee Comprehensive Assessment Program (TCAP), a set of standardized tests taken by all students in grades 3-8 each April. This TVAAS rating was used to calculate teachers' and principals' overall effectiveness ratings.

Research Questions

The data collection for the study was guided by the following research question:

- What theoretical model can explain how teachers have adapted their practices in response to value-added evaluation policy?

From this central question, two empirical questions followed:

- What are the lived experiences of teachers who have been evaluated by the Tennessee Value-Added Assessment System?
- What practices do teachers choose in response to Tennessee Value-Added Assessment System evaluation policy?

Theoretical and Conceptual Frameworks

Within the last twenty years, Coburn (2001), Marz, Kelchtermans, and Dumay, (2016), and others have established a small but growing body of work that continues to

draw from institutional policy and sensemaking to explore the process of education policy implementation. Kim and Youngs (2016) suggested that teachers and other actors within schools incorporate institutional values and norms into their own practices based on their sensemaking of the goals or tasks the policy dictates, their organizational contexts and their professional identity. To illuminate the implementation of VAEP policy, researchers need to review the interactive dynamics among the institutional values and beliefs embedded within VAEP, how teachers respond to those values through meaning making and practice, and how those responses in turn shape institutional structures (Hallett & Meanwell, 2016). Integrating theories of institutional theory, sensemaking, and agency provides an inclusive framework for exploring these recursive relationships.

Sociocultural frames recognize that cultural, historical, and social structures reflected in “mediational tools ... such as policy mandates, curriculum guidelines, and state standards” (Lasky, 2005, p. 900) influence what individuals believe, what they think, and how they act. However, frameworks to understand how public policy is implemented through teachers’ practices within their social environments are limited (Golden, McLeroy, Green, Earp, & Lieberman, 2015). Teachers’ perceptions and adaptations to value-added evaluation have been well-documented, but few studies offer culturally-situated and empirically-based explanations for teachers’ responses. Given the paramount role teachers play in schools and in students’ lives and the wide-spread implementation of VAEP in public schools, it is surprising that little research has been conducted to understand the process of how teachers make sense of policy and choose

practices when responding to this policy. Drawing from the work of Stokols (1992), McLeroy, Bibeau, Steckler, and Glanz (1988), and Golden et al. (2015) on social ecologies and public policy, this study integrates theories of institutional theory, sensemaking and human agency to provide a conceptual framework for understanding teachers' responses to policy within micro and macro-level contexts.

Summary of Methods and Analysis

The data were collected from single, semi-structured interviews with 19 elementary education teachers in Tennessee, a state that implemented a value-added component to teacher evaluations in 2011. Interviews were audio recorded and transcribed for coding, and a process framework was used for analysis. A theoretical model was constructed from the self-reported beliefs, perceptions, and experiences of the participants. The model addressed the central question by describing and relating:

- the structural conditions that underlie the development of teachers' adaptations in practice, including institutional beliefs and values,
- the phenomena that arose from the structural conditions,
- the contexts and components of personal histories that influence teachers' choices of practice,
- teachers' responses to policy implementation through their adaptive practices, and
- the consequences of teachers' adaptive practices, and how in turn, those practices shape institutional structures.

Biases and Assumptions

Critical Reference Group Membership

Genat (2009) defined the critical group of participants in social research as “the particular stakeholder group whose experience and knowledge is unknown or perhaps subjugated” (p. 105). Wadsworth (2011) called these stakeholders the “critical reference group” (p. 21). This study embraces a critical reference group approach that identifies with the interests of the critical reference group, has a deep respect for its members and recognizes the legitimacy of their emotions, beliefs, values, and perspectives (Wadsworth, 2011).

Teachers who have been evaluated with value-added methods as employees of Community School District (a pseudonym) were recruited to participate as the critical reference group for this study. As an elementary school teacher for the past seven years within Community School District, the researcher has extensive direct and personal experiences with the critical reference group’s situation. In addition, many of the participants were past or current colleagues of the researcher, and all of the participants had met the researcher in a professional capacity prior to being invited to join the study. The researcher’s position as a fellow teacher and colleague of the critical reference group facilitated trust and candidness in the interviews, but it also created conditions and limitations of bias. To mitigate the effects of the researcher’s standpoint, the researcher did not reference prior knowledge and experiences within Community School District during interviews. If a participant commented on the researcher’s presence or knowledge of a practice or narrative, the researcher specifically requested that the participant explain

his or her experience in full detail. To diminish the influence of the researcher's prior experiences in the analysis of teachers' interview responses, the coding process relied on the words and narratives of the teachers in the study instead of preconceived categories.

Emancipatory Purpose of Education

“All educational practice implies a theoretical stance on the part of the educator” (Freire, 1970a, p. 6). There is no such thing as a neutral education because it either serves to conform learners to the current paradigm or it becomes a “practice of freedom” that encourages learners to engage with and transform their reality (Freire, 1970b, p. 34). This research embraces Freire's emancipatory approach to education based on the assumption that people can engage with the paradigms of institutions to make changes within them; the study also challenges practitioners, administrators, and policy-makers in public education to answer Freire's question, “on what side are we when we teach/act?” (Mayo, 2003). An emancipatory education encourages the questioning of society's structures and its norms and values. However, Freire's conception of conscientisation, or problematizing societal structures, cannot happen within the “banking” model of education (Ruggunan & Spiller, 2014). This traditional pedagogical approach positions teachers as experts who determine what constitutes knowledge and “make deposits” that students passively receive without critical reflection or inquiry (Freire, 1970b). bell hooks (2003) warned that these conventional pedagogies propagate “obedience to authority and accepting of dominator-based hierarchy” (p. 19-20). Although the study espoused a particular approach to education, the researcher intentionally presented a balanced literature review of education policy.

Definition of Terms

Policy Paradigms

- **Educational Accountability:** The concept of holding schools, districts, educators, and students responsible for educational outcomes (Editorial Projects in Education Research Center, 2004). “New” accountability systems within the current standards-based reform focus on performance, regard schools and teachers as units of improvement, include continuous improvement strategies, inspections, public reporting, and consequences attached to performance levels (Fuhrman, 1999). Reeves (2005) posited that this kind of results-driven accountability values student test scores as the most accurate measure of teacher quality.
- **Neoliberalism:** A collection of values, ideologies, policies and practices that promote the globalization of free markets through expansion in underdeveloped countries, privatization of public services, and deregulation of private industry. Neoliberal thought decentralizes decision-making to the individual and shifts social responsibility from the community to the individual (Gibson & Ross, 2007; Turner, 2008).
- **Free Market:** A idealized system in which prices are based on competition through unregulated economic exchanges and centralized interventions by government (taxes, quotas, tariffs, etc.) are minimal or non-existent (Merriam-Webster, n.d.; Orlitzky, 2016). Through the invisible-hand

mechanism of self-regulating behavior, society benefits by having self-interested actors make free economic decisions that benefit them (Orlitzky, 2016).

- **Scientific Management:** A theory of management in factory production in which management analyzes the production process using “scientific” methods to determine the most efficient ways for workers to complete specialized tasks, thereby placing control over the production process with management (Au, 2011).

Value-Added Methods and Policy

- **Claim Students:** A roster verification process used to link individual students to a teacher and include those students in the statistical analysis used to calculate a teacher’s TVAAS scores (SAS Institute, 2017).
- **Student Growth:** A student’s current achievement compared to all prior achievement, with achievement being measured by a standardized assessment (SAS Institute, 2017).
- **Tennessee Value-Added Assessment System (TVAAS):** A statistical analysis that measures student growth year over year regardless of the student’s proficiency on the state assessment (Tennessee Department of Education, n.d.).
- **TVAAS Score:** A comparative measure of a student’s achievement on the state’s standardized test relative to the performance of his or her peers who

performed similarly on past assessments (Tennessee Department of Education, n.d.).

- **Tested Grades:** The elementary school grades required to sit for the state standardized tests each spring, specifically grades three through five in Tennessee public schools.
- **Validity:** The extent to which a tool accurately measures the construct it was designed to assess.
- **Value-Added Evaluation Policies (VAEP):** policies and evaluation systems that incorporate measures of student growth on standardized state tests as a factor in evaluating teachers, specifically for calculating teachers' effectiveness scores.
- **Value-Added Methods or Models:** Statistical methods or models designed to measure a teacher's impact on student achievement; the value he or she adds, apart from other factors that affect achievement, such as individual ability, family environment, past schooling, and the influence of peers (RAND Corporation, 2002).
- **Value-Added Rating or Score:** In Tennessee, a teacher rating of one to five based on the growth scores of his or her claimed students. To earn an individual growth score (instead of a school wide score), teachers must be able to claim at least six full-time equivalent (FTE) students.

Teaching and Learning

- **Departmentalization:** An arrangement in which a teacher will teach one or two subject areas for two or more classes of students in a daily or weekly schedule.
- **Instructional Time:** The time students spend in schools devoted to academic activities and tasks.
- **Leveling:** The practice of grouping students into self-contained classrooms or traveling homerooms (if the grade level is departmentalized) based on their abilities.
- **Practice:** The application or use of an idea, belief, or method of teaching and learning as opposed to theories about such application or use.
- **Standards:** a common set of expectations for the knowledge students will gain and the tasks students will be able to perform at the end of a grade for each subject area.

Since the passage of NCLB in 2001, teachers have emerged as the primary focus of education policy. Accountability is now the dominant paradigm for improving K-12 education, resulting in evaluation policies designed to directly impact teacher behavior using a competitive system with high-stakes consequences. Analyzing teachers' understanding of policy through their reflections and narratives is vital for illuminating the impact of teachers' sensemaking on policy implementation. In addition, locating teachers' responses to policy within multiple organizational ecologies is essential for analyzing and improving education policy implementation. This study explores the

beliefs, opinions, and experiences of elementary public-school teachers who have responded to state-mandated value-added evaluation policies.

Organization of the Dissertation

This dissertation is organized into five chapters. Chapter 1 begins with an introduction to the study and provides the purpose of the research project. The chapter also briefly outlines the context of the study. Chapter 2 presents the assumptions and paradigms of the study within discussions of the conceptual framework and foundational theories. A literature review of value-added evaluation policies is followed by a genealogy of value-added methods in teacher evaluation policy. The chapter concludes with a review and critique of research on the outcomes of value-added evaluation policies, including teachers' perceptions and responses in practice. The research design for this study is described in detail in Chapter 3, which also provides the rationale for using a qualitative, grounded-theory methodology. Chapter 4 presents the study's findings, which include the core themes generated from the open and selective codes and the theoretical model generated to respond to the central research question. Chapter 5 provides a response to the empirical research questions and situates the theoretical model within grounded theory reliability criteria, within the conceptual framework presented in the literature review, and within value-added evaluation policy literature discussed in Chapter 2. Discussions of the implications for practice and policy as well as recommendations for further research are also included. The chapter concludes with the researcher's reflections and closing remarks.

CHAPTER II

REVIEW OF LITERATURE

Review of Study

The purpose of this qualitative grounded theory study is to construct a theoretical model that explains the process teachers utilize to choose their practices in response to a state-mandated, value-added teacher evaluation policy.

The data collection for the study is guided by the research question:

- What theoretical model can explain how teachers have adapted their practices in response to value-added evaluation?

From this central question, two empirical questions follow:

- What are the lived experiences of teachers who have been evaluated by the Tennessee Value-Added Assessment System?
- What practices do teachers choose in response to Tennessee Value-Added Assessment System evaluation policy?

Preview of Chapter

This study considers the use of value-added scores in teacher evaluations in Tennessee as a critical case to explore the relationships between sociocultural beliefs embedded in teacher evaluation policy, how teachers understand value-added evaluation policy, and how their understanding of policy shapes their practices in classrooms and

schools. The chapter begins with a discussion of the conceptual framework and related foundational theories to present the assumptions and paradigms of the study. The second section of the chapter includes a literature review of value-added evaluation policy beginning with macro-level sources of the ideologies embedded in value-added evaluation systems. The section continues with a genealogy of value-added methods in teacher evaluation policy. The chapter concludes with a review of research on the outcomes of value-added evaluation policy, including teachers' perceptions and responses in practice and the impact of teachers' adaptations.

Conceptual Framework and Theoretical Foundations

Conceptual Framework: Sociocultural and Ecological Models

Jiang et al. and Tuytens and Devos (2015; 2009) suggested that teacher perceptions profoundly affect outcomes of policy. However, these perceptions are not formed in a vacuum; responses of teachers to policy are informed by social constructions of self-image, school and community discourse, public narratives, and larger political and cultural ideologies (Burger, 1987; Day, 2002; Flores & Day, 2006). Similarly, Braun (2015) suggested a framework for understanding teachers' adaptations that accounts for the complex dynamics sparked by a new evaluation policy; political and educational contexts influence program design, which impacts behavioral responses of teachers with long-term effects. Braun stated, "Policymakers should move toward a more ecological approach to the design of accountability systems" (2015, p. 130). To explore how teachers have responded to value-added evaluation through practice, this study employs a

social-ecological framework adapted from ecological models of human development and health behavior. The framework is intended to illuminate how multiple social contexts interact in the policy implementation process.

Tabàra and Chabay (2013) suggested that “knowledge is mostly what works in a particular context” (p. 75). Social ecological models focus on people’s transactions with their sociocultural surroundings, or “environments” in contrast to behavioral models and theories that do not consider broader community, organizational, and policy influences on behaviors (Sallis & Owen, 2015). Within systems theory and social ecology, Stokol’s (1992) proposed “cycles of mutual influence” (p. 8) between people and their social settings offer another explanation of how policy and practice shape each other. Similarly, Bronfenbrenner’s (1979) social ecological model of human development examines people’s transactions with their physical and sociocultural surroundings or *environments* to understand behavior (Golden & Earp, 2012).

While *ecological* is a term derived from the biological and environmental sciences, here it refers to relationships between teachers and their social environments. Social-ecological systems (SES) recognize multiple factors affecting specific behaviors, including influences within an individual and his or her social environments (interpersonal, organizational, community, policy, and ideological, or super-structural relationships); in addition, these influences can interact and cause changes across all levels (Sallis & Owen, 2015). Therefore, the social-ecological model used in this study assumes that teachers affect their setting, and the changed setting influences teacher behaviors. While the extant literature on SES is largely applied to developing and

evaluating public health behaviors and policy, this study drew upon SES to develop a conceptual framework for analyzing teachers' responses to an evaluation policy intended to influence teacher practices. Golden et al. (2015) described socio-ecological models as "visual depictions of the dynamic relationships among individuals, groups, and their environments" (p. 10S). Figure 1 is an adaptation of the McLeroy et al. (1988) socio-ecological model that uses concentric circles to illustrate how contexts are interconnected. It is not intended to represent a fixed or closed policy process or an ahistorical model for producing knowledge about evaluation policy and teacher behavior. Instead, it attempts to facilitate an understanding of the impact of evaluation policy embedded within a specific school organization, operating under a specific district policy as directed by state and federal law, within a specific national and global historical moment. Socio-ecological models for understanding the dynamics of public policy suggest that multiple levels and sociological environments mutually influence each other. Therefore, evaluation policy intended to influence individual teacher behavior is mediated by public discourse, organizational culture, interpersonal relationships between teachers and students, as well as individual teacher's beliefs. In addition, all of these environments are situated within a superstructure, a context of ideologies and paradigms that shape society's beliefs about education reform (see Figure 1).

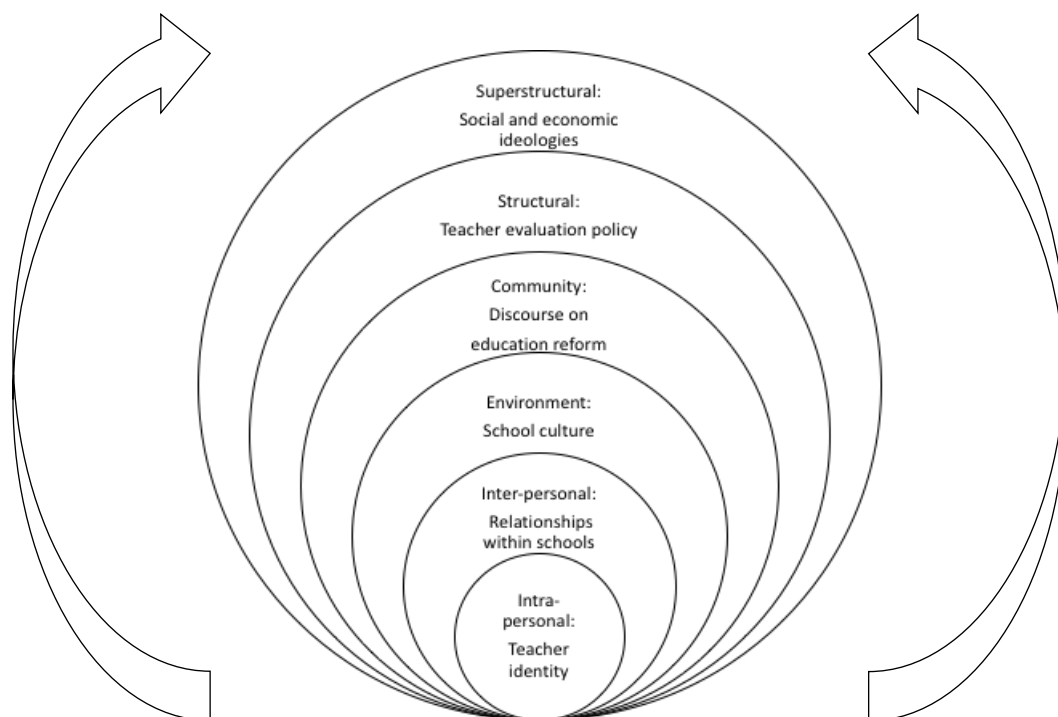


Figure 1. Socio-ecological model adapted from McLeroy et al, 1988.

Socio-ecological approaches to understanding teacher responses to reform recognize that what individuals believe and think, and how they act is influenced by cultural, historical, and social structures reflected in formal policies, guidelines and standards (Lasky, 2005). Socio-ecological models used for understanding policy implementation are “complex adaptive systems, continuously evolving and adapting at different scales and levels of an organization” (Anderies & Janssen, 2013, p. 514). Social-ecological models (SEM) differ from traditional behavioral models and theories that focus on an individual teacher’s personality, skills, and the social environment of his

or her classroom (McLeroy et al., 1988; Sallis & Owen, 2015). In contrast, SEM account for broader influences including school culture, state and federal policy, public discourse on education, and the global socio-political climate. Proponents of SEM claimed that the current ideology surrounding individual responsibility limits policymakers' understanding of the dynamics and complexities of social systems, while the dominant discourse perpetuates those beliefs about what shapes behavior (Golden et al., 2015; McLeroy et al., 1988). McLeroy et al. (1988) suggested a broad, multi-faceted approach to promoting improved public health behaviors that takes into account the social context of individual behavior. The authors disputed "the ideology of individual responsibility that ignores what is known about human behavior" (1988, p. 352). Like other free market-based reforms, value-added evaluation is a policy tool for states and districts to change the behaviors of individuals rather than change the conditions and environments in which the policies are formed. Social-ecological models challenge the ideology of individual responsibility that belies the complexities of social causation (McLeroy et al., 1988, p. 352).

Institutional Theory and Policy Implementation

Policy implementation is a daily event in schools and districts" (Rigby, Woulfin, & Marz, 2016, p. 295). As in all institutions, public education organizations reside within multiple, nested levels of social contexts that interact and shape one another. To analyze and understand education policy implementation, researchers must conceptualize and integrate ideologies of the macro level superstructure, the daily activities of teachers within schools, and the many social environments layered between them. While teachers

exercise agency within their schools, the communities, districts, agencies, and governments at the state and federal level and the broader global policy environment also enable and constrain their practice through institutional structures.

Although rarely implemented as it is intended or written, policy shapes the daily work of teachers (Rigby et al., 2016, p. 295). Coburn (2016) described policy as “a set of rules, often supported by resources, that attempts to constrain or channel behavior in particular directions through regulative, normative, or cognitive means” (p. 466). While formal evaluation policy is a regulative tool for institutions and organizations to exert pressure on teacher behaviors, the institutional practices, assumptions, values, and rules embedded in policy, called institutional logics, also influence teachers’ responses to policy (Woulfin, 2016; Thornton & Ocasio, 2008).

Hallett and Meanwell (2016) suggested that policy paradigms, a form of institutional logics, shape policy by providing frames for identifying problems and by limiting the range of possible solutions. These logics are socially constructed and reflect the dominant ideologies of a particular context. Therefore, teachers’ understandings of policy are shaped by larger sociocultural paradigms and logics that enable and limit social interactions within the school organization. Policy does not institutionalize itself; it requires the agency of individuals to act on interpretations of policy logics (März et al., 2016). Honig (2006) noted that only since the 1990s has research in education policy explored how the agency of teachers and other school-professionals shapes implementation. Furthermore, “The institutional tradition needs to pay more attention to human agency in its efforts to understand relations between policy and the technical core

of institutions” (Spillane & Burch, 2006, p. 95). Implementation depends upon the ability of a policy to achieve specific objectives and desired goals by shaping individual and collective action (Coburn, 2016). Rigby et al. (2016) suggested that shifts in institutional logics at the macro level offer opportunities to see how local policies are implemented and institutionalized through local actors’ agency at the micro level. Therefore, education policy research will benefit from analyzing the dynamics between institutional structures and teacher agency to illuminate how policy is institutionalized in schools.

This study draws on institutional theory to analyze the dynamics between teacher evaluation policy and teachers’ practices. Institutional theory also provides a lens to look closely at how teacher agency is constructed, exercised, and constrained by organizational structures and practices. In particular, it employs a lens to analyze how teachers’ practices are shaped through changes in policy. Institutional theory is useful for exploring how institutionalized structures dictated by policy shape how teachers practice in schools (Bray & Russell, 2016, p. 369).

Institutional theory examines the processes by which institutions establish social behavior through shared beliefs, norms, rules and practices, called institutional structures. Giddens defines institutional structures as the ways we understand what behaviors are expected, how they should be performed and the practices and resources that support those understandings (Rigby et al., 2016, p. 296). Within a social institution, structures create rules of behavior, benefits, and punishments for its members (Ratner, 2000). While early institutionalism of the nineteenth century focused on how institutions account for social behavior, by the 1970s a new iteration of institutional theory, neo-institutionalism,

had emerged (Scott, 2005; Thornton & Ocasio, 2008). DiMaggio and Powell (1991) described how “new institutionalism ... comprises a rejection of rational-actor models, an interest in institutions as independent variables, a turn toward cognitive and cultural explanations, and an interest in properties of supraindividual units of analysis that cannot be reduced to aggregations or direct consequences of individuals' attributes or motives” (p. 8).

Despite its socio-cultural turn, new institutional theory, known as neo-institutionalism, has been criticized in recent decades. Hallett (2010) contended that people, their work activities, social interactions, and sensemaking processes tend to be overlooked by the macro level focus of contemporary neo-institutionalism. Coburn (2001) pointed out that early neo-institutional theory tended to ignore localized responses to institutional logics. Hallett and Meanwell stressed that neo-institutional theory has misconceived institutional logics as unilateral pressures on behavior and has underestimated teacher agency in understanding how teachers' adapt to reform (2016; 2010).

Within neo-institutional theory, several attempts have been made to reconcile the structure-agency debate or the tension between actors' ability to exercise agency within institutions and the pressures exerted by institutions which enable and constrain actors' behaviors (Coburn, 2016; Hallett & Meanwell, 2016; März et al., 2016; Giddens, 1979; Woulfin, 2016). Giddens reasoned that the structure-agency tension is a false binary because action and structures are both socially constructed and (re)create one another (K. H. Tucker, 1998). Giddens' (1979) structuration theory attempted to reconcile the

structure-agency conflict by linking macro level institutional pressures with micro level action; with structuring, institutions shape actors, who in turn reaffirm or modify structures through their actions. Sometimes referred to as “discursive institutionalism, ” inhabited institutionalism recognizes the constraints of institutional logics on behavior but also asserts that those logics are mediated, enacted, interpreted, and negotiated by actors (Hallett & Meanwell, 2016; März et al., 2016). März et al. (2016) suggested that inhabited institutionalism looks at how teachers “talk back” to pressure from institutional logics of value-added evaluation. Through these dynamic interactions, the behaviors and practices of actors within organizations are linked to larger societal structures. Woulfin (2016) also added to the scholarship on neo-institutionalism with her concept of “lived logics” (p. 338) which looked at the ways institutional logics are acted out in the practices of organizational actors.

Rejecting rational choice theory, which looks to individual choices and preferences to explain social behavior, Spillane and Burch (2006) asserted that “understanding policy implementation requires more than the assumption that teachers understand what policy is asking them to do and that they should simply adopt, ignore, or modify policy guidance” (p. 95). Kelchtermans et al. (2016) encouraged an approach to the study of teachers’ responses to policy that advocates sensemaking, a process based on a personal system of knowledge and beliefs, to explain how teachers consider their professional situation, give meaning to it, and subsequently act within it. Spillane & Burch (2006) supported this assertion: “Because teachers use their prior knowledge and experience to make sense of the ideas pressed by policy, policy to practice connections

are mediated by teacher sense making” (p. 95). Teachers and other actors within schools interpret policy logics and incorporate them into their practices based on their “sense” about the goals or tasks the policy dictates, their organizational contexts, and their professional identities (Kim & Youngs, 2016). Coburn (2001) suggested that sensemaking theory provides a lens to explore how teachers “mediate norms, belief systems, and practices that have diffused from the institutional environment” and reconstruct them as their own (p. 147). Schmidt and Datnow (2005) found that teachers made sense of reforms through their own classroom practices. Sensemaking theory brings balance to the structure-agency tension within neo-institutionalism by accounting for the agency that actors exercise within institutional structures.

Education reforms of the early 21st century have focused on increasing teacher accountability through teacher evaluation policy. Evaluation policies communicate outcome expectations, and teachers are rewarded or sanctioned depending on how closely their performance matches the expectations. Most research has focused on the results of teacher evaluation policy implementation while relatively few studies have analyzed teacher evaluation policy from a sensemaking perspective (Kim & Youngs, 2016). Kim and Youngs (2016) claimed “Teachers and principals are likely to actively interpret and respond to the policies rather than just follow their directions” (p. 2). Value-added evaluation was designed to produce outcomes, not to dictate an instructional program or approach; the policy does not provide directives for achieving the desired outcomes. Spillane and Burch (2006) suggested “that when policies are less elaborated, teachers’ sense making produces qualitatively different understanding among teachers” (p. 95).

Therefore, sensemaking theories are necessary for understanding teachers' responses and their implementation of evaluation policy.

Institutional theories explain how structures disseminate institutional beliefs, values, and norms through logics, but more is needed to explain how actors in organizations "talk back," (März et al., 2016, p. 308). Analyzing actors' sensemaking and interactions with a changing institutional environment can provide detailed evidence illustrating how organizations are places where people and groups make sense of and creatively use institutional logics (Rigby et al., 2016, p. 299). Sensemaking theory draws from Blumer's (1969) theory of symbolic interaction for a heuristic method of understanding how teachers react to reform based on the meanings the reforms have for them (Hallett & Meanwell, 2016). Kelchtermans (2009) emphasized that a personal interpretive framework comprised of two interconnected domains, professional self-understanding and subjective educational theory, operates as a sensemaking filter for teachers to observe, interpret, and evaluate calls for change and innovation. Table 1 describes the five components of professional self-understanding, a construct for describing teachers' representations of themselves that develops over time through social interaction (Kelchtermans, 2009).

Table 1. Kelchtermans' components and descriptions of professional self-understanding

Component of Professional Self-Understanding	Description
Self-image	How I describe myself and how I think others describe me too
Self-esteem	How well I think I am doing my job
Task Perception	What I think I should be doing
Job Motivation	What makes me want to stay in this job (or leave it)
Future Perspective	What I expect about my future in this job

Adapted from Kelchtermans (2009).

Subjective educational theory suggested that teachers use a personal system of knowledge and beliefs about education that develops over the course of their careers. Sources for this system stem from training and formal and informal education. By exploring this personal system of knowledge and beliefs, researchers can conceptualize how teachers make sense of reforms (Kelchtermans, 2009). The personal interpretive framework is a cognitive and affective lens through which teachers “look at their job situation, give meaning to it, and act in it.”(März et al., 2016, p. 309). Teachers make sense of policy through an interpretive lens based on their personal experiences and knowledge thereby shaping their perceptions of agency in responding to policy. Therefore, conceptualizations of human agency from social cognitive theory are helpful for understanding how sensemaking and cognition relate to action.

Theories of Human Agency

Agency, the capacity to act, choose, and imagine, is central to the human condition (Brockmeier, 2009). Bandura's (2001) conception of human agency within

social cognitive theory rejected behavioristic views that human behavior is shaped entirely by environmental stimuli. He reasoned that because judgments and actions are partly self-determined, people can effect change in themselves and their situations through their own efforts (Bandura, 1989). Agentic actors reflect on their world, decide when to act upon it, and then reflect on their interventions (Bhaskar, 1989).

Human agency enables people to play a part in their self-development, adaptation, and self-renewal with changing times (Bandura, 2001). Agency provides a theory of action to help explain individuals' behavior based on their sensemaking and is especially helpful for understanding responses to evaluation policy. Agency is exercised through the meaning actors make of their environment. Emirbayer and Mische (1998) identified the "practical-evaluative" dimension of agency as the capacity of actors to make practical and normative judgments among alternative possible paths of action "in response to the emerging demands, dilemmas, and ambiguities of presently evolving situations" (p. 971). In exploring responses of teachers to value-added evaluation policy, agency is a useful construct for analyzing action based on sensemaking of changes in the institutional environment. Bandura's social cognitive theory of agency provides a bridge between sensemaking and praxis.

Self-efficacy is the degree to which an individual believes that he or she is capable of successfully performing a specific behavior (Prussia & Kinicki, 1996). The belief of self-efficacy is the foundation of human agency and the most pervasive mechanism for exercising it (Bandura, 1989). It affects thoughts, motivation, and life choices and enables people to predict the occurrence of events and to create the means for

exercising control over those that affect their daily lives (Bandura, 1989). Self-efficacy beliefs are largely shaped by the results from an individual's prior attempts at mastering a task (Driscoll, 2005). An individual's interpretations of success or failure at accomplishing the task, his or her perceptions of the difficulty of the task, and how much effort is required mediate how those experiences affect self-efficacy beliefs (Driscoll, 2005). In turn, Bandura and Jourden's (1991) research suggested that perceived self-efficacy is a significant determinant in mastering a performance task. Therefore, beliefs in self-efficacy and performance experiences are interdependent.

When people do not have direct control over their social conditions and institutional practices, they may pursue their well-being and desired outcomes by exercising proxy agency (Bandura, 2001). However, for the purposes of this study, Bandura's conception of direct personal agency is most useful for understanding how teachers respond to policy. The discussion of direct personal agency focuses on three characteristics consistent with foundational institutional theories of this study; agency is conceptualized as transactional, socially situated and interactive. Agency does not reside within structures or actors. Bandura (2001) insisted that it is expressed through relationships with others and with social structures rather than residing as a discrete entity in a particular place. Furthermore, in agentic transactions, people are creators as well as creations of social systems (Bandura, 2001). Karl Marx (1976) suggested, "but the essence of man is no abstraction inherent in each single individual. In its reality it is the ensemble of the social relations" (vol. 5). Agency is not a transcendental capacity but a "socialized body, investing in its practice socially constructed organizing principles that

are acquired in the course of a situated and dated social experience” (Bourdieu, 2000, p. 137). However, agency is also conceptualized as the ability of individuals and groups to resist structural forces by exercising their will in challenging those structures through day-to-day practices (Basu & Dutta, 2008). Marcuse compared agency to justice, morality, intelligence, sensitivity, and language; none exist *a priori* or as an intrinsic quality of the individual; instead, they must be realized through social reforms (Ratner, 2000). For the purposes of this study, agency refers to the ability of individuals to make sense of their environment, choose, and commit to a course of action meaningful in their particular context (Mitra, 2012). To unify the characteristics of agency as transactional and socially situated, social cognitive theory offers a model of triadic reciprocal causation to illustrate how personal factors, behavior, and an individual’s external environment interact to direct human action (Bandura, 1989; Wood & Bandura, 1989). Through agency, people exert personal influence to develop, adapt, and renew themselves as their environment changes (Bandura, 2001). Agency is a product of a culture as much as the individual and social and cultural conditions shape broad conceptions of agency within a society (Ratner, 2000). How human agency and adaptation are understood and analyzed shapes policy responses to social challenges. Cultural ideologies regarding agency become part of institutional structures and shape policy design. For example, policies that emphasize individual responsibility and an individualistic approach to agency measure success in terms of economic contribution rather than well-being (Brown & Westaway, 2011). As a result, individual responsibility becomes embedded as an institutional logic through policy.

Given the role of society and culture in shaping agency, Emirbayer and Mische (1998) cautioned that conceptions of agency as a social construction are often “so tightly bound to structure” that it is difficult to see how agency actually shapes social action (p. 963). Bourdieu’s (2000) sociology of culture employs the concept of habitus to reconcile individual agency while still maintaining that this power to construct social reality is itself socially constructed. Bourdieu’s conception of habitus is composed of our thoughts, tastes, beliefs, values, interests, and our understanding of the world constructed through social relations with family, culture, and education (Reed-Danahay, 2005). Therefore, conceptions of agency are not universal but developed within cultures and framed by habitus. An individual exercises agency through the habitus, yet that individual’s understanding of agency depends upon his or her experience with social conditions (Ratner, 2000). Likewise, an individual’s sense of agency and his or her ability to act cannot be separated from the effects that policies have on shaping him or her (Lasky, 2005). From a cultural perspective, agency “both reproduces and transforms [environmental] structures in interactive responses to the problems posed by changing historical situations” (Emirbayer & Mische, 1998, p. 970).

This study integrated institutional, sensemaking, and agency theories to develop a conceptual framework illustrated in Figures 2 and 3 that recognizes both the constraints of institutional structures and the agency of individuals and groups within institutions.

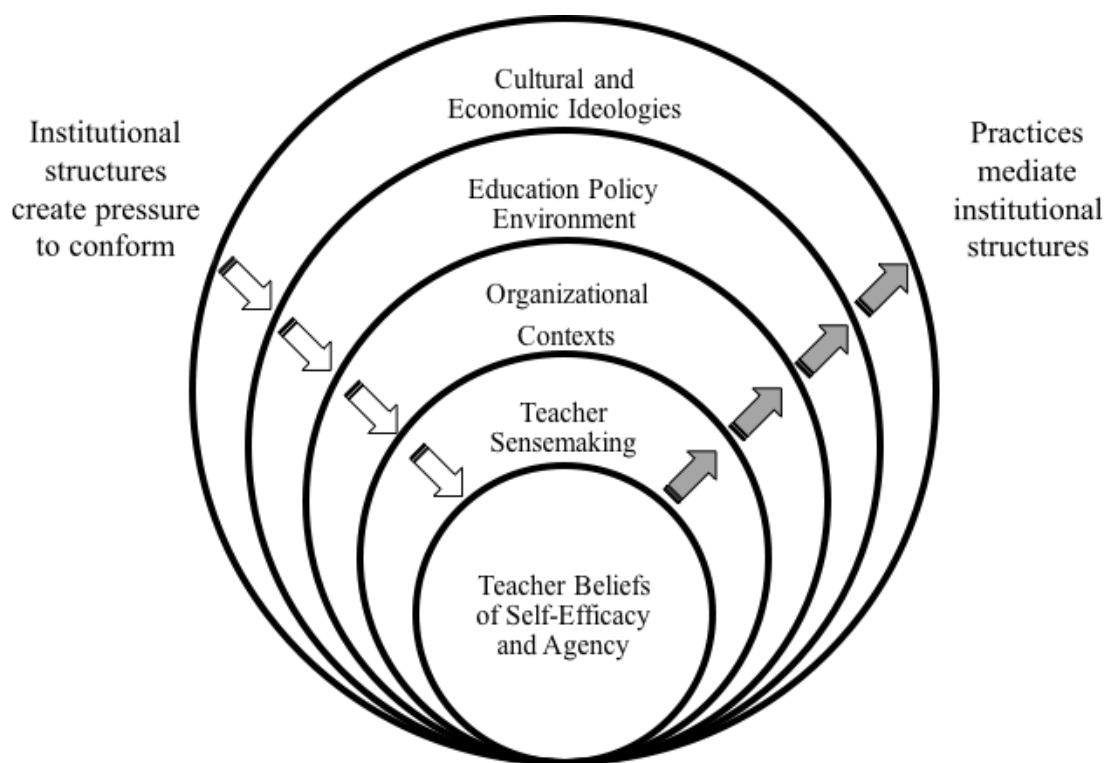


Figure 2. Conceptual framework integrating institutional, sensemaking, and agency theories.

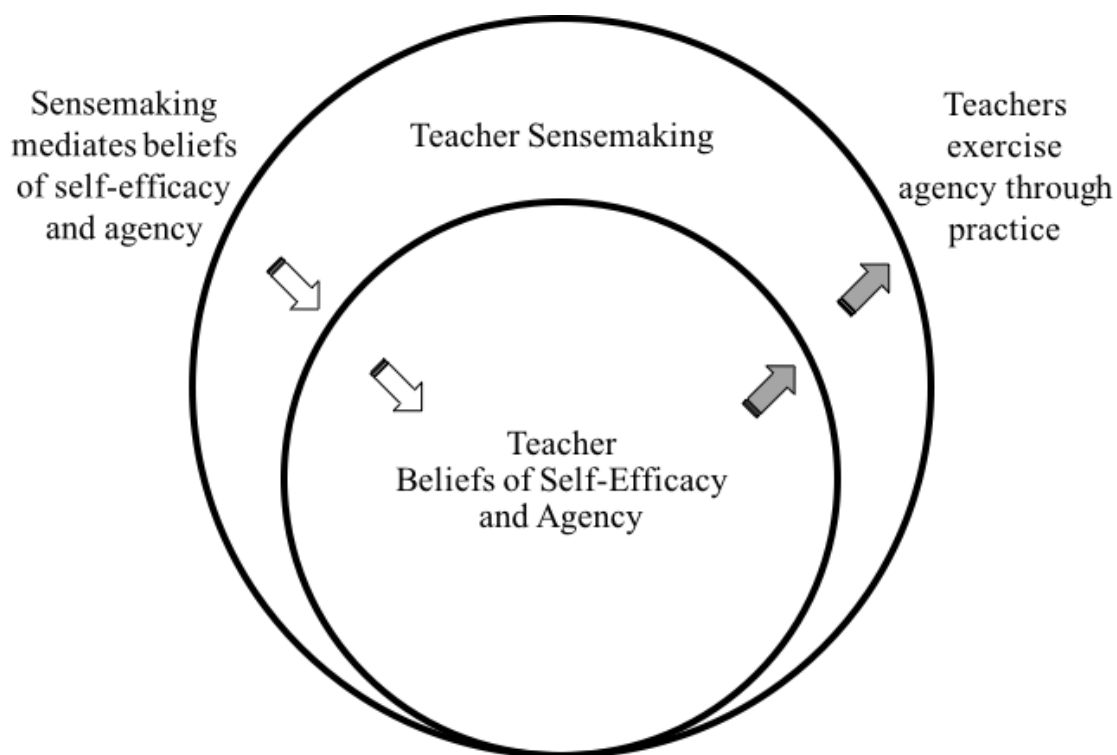


Figure 3. Detail of conceptual framework integrating institutional, sensemaking, and agency theories.

Institutional theories offer explanations of how socially and culturally constructed structures within organizations create pressure to conform and constrain people's behaviors. However, these theories alone are insufficient for explaining how actors exercise agency within institutions and organizations. Sensemaking theory is useful for explaining how actors make meaning of their environments to decide how to respond to challenges or changes. Theories of agency describe the environmental and internal factors that determine human action. Therefore, theories of sensemaking and agency are necessary to understand teachers' socially-situated responses to institutional pressures of policy and logics. Given its complexity of actors, structures, and stakeholders, education

policy implementation is a useful topic to explore the interaction between structure and agency (Rigby et al., 2016). Research illuminating these dynamics can add to both the empirical and theoretical knowledge base of the structure-agency debate.

Dimensions of Teaching that Impact Sensemaking

Three dimensions inherent in teaching shape how teachers make sense of high-stakes, value-added evaluation policy:

- Teaching as a humanistic vocation
- Structural vulnerability
- Relational work

Humanistic vocation. Dewey (1916) described vocations as being one's direction in life. For many educators, teaching is a vocation, which comes from *vocare*, the Latin word for "to call." Buechner (1993) depicted a vocation as "the place God calls you to is the place where your deep gladness and the world's deep hunger meet" (p. 18-19). Bill Ayers (2001) described teaching as "the vocation of vocations because to choose teaching is to choose to enable the choices of others" (p. 24). According to Fried, (1995) a passionate teacher is "someone in love with a field of knowledge, deeply stirred by issues and ideas that challenge our world, drawn to the dilemmas and potentials of the young people who come into class each day" (p. 1). Many teachers identify their work as more than a profession but a calling or vocation. In their study of elementary school teachers, Woods and Jeffrey (2002) found that teachers have a "strong emotional dedication ... moral and political investment ... and commitment" to their work (p, 93). However the intrinsic, non-instrumental view of education has been replaced with an emphasis on

more functional and results-driven forms of teaching through results-based accountability reforms (Day, 2002). Teachers have responded by adapting their identities to meet the demands of audits and performances, which has placed them at odds with the humanistic vocationalism of their professional lives (Woods & Jeffrey, 2002).

Structural vulnerability. Kelchtermans (2005) suggested that vulnerability is a structural condition of teaching that arises from the ethical dimension of teachers' interactions and teaching practices. To teach is to be vulnerable; it is the way in which "teachers live in their job situation" (Kelchtermans, 1996, p. 307). Palmer (1998) reflected that teaching is a "daily exercise in vulnerability" (p. 17). Kelchtermans (2009) insisted that vulnerability in teaching is not understood as a feeling or experience of emotion but, in fact, as structural condition of the profession.

In their study of vulnerability in teacher peer groups, Uitto et al. (2016) presented empirical evidence of Kelchtermans' (2009) three sources of vulnerability in teaching: the inability to control essential working conditions, the difficulty to prove one's effectiveness, and the inevitable uncertainty in their decision making that has moral consequences. Vulnerability arises from the ethical dimension of teachers' interactions and teaching practices. van Veen et al. (2005) contended that the teaching relationship is not just technical or intentional; while the instrumental outcomes of teaching are a concern, there is more "at stake" (p. 918). Teachers rely upon their values to make decisions that have moral consequences because they affect the lives and needs of children (Kelchtermans, 2005). Uitto et al. (2016) explained that the moral dimension of teaching stems from the daily, ongoing "process of value-laden decision making about

how to do justice to pupils' educational needs" (p. 7-8). At the same time, a teacher can never have full command of the outcomes of his or her actions because the pedagogical relationship "radically escapes control and intervention" (Fenstermacher, 1990, p. 998). The vulnerability inherent in teaching is experienced by teachers as a sense that professional identity and moral integrity are always in question (Kelchtermans, 1996). This relationship of moral responsibility combined with teachers' lack of control over the outcomes of teaching creates conditions of ethical vulnerability.

Palmer (1998) reflected that "unlike many professions, teaching is always done at the dangerous intersection of personal and public life (p. 17). van Veen et al. (2005) suggested several reasons why value-added evaluation policies may amplify teacher perceptions of professional vulnerability. van Veen et al. (2005) claimed that a teacher's identity is at risk when confronted with reforms because of the conditions of ethical vulnerability in teaching. Johnson (2015) contended that when value-added ratings in teacher evaluations are used to inform decisions about compensation, promotion, and retention, they raise the stakes even higher for teachers. Ball (2003) suggested that as a result of new monitoring and accountability systems, education reform has evoked high levels of uncertainty, instability, and vulnerability among teachers. Teachers have little or no control over the evaluation policies that determine their job effectiveness, leaving them with little agency in how they are evaluated. Finally, because of the ethical responsibilities of teaching, a poor job evaluation affects teachers' moral dimension of self-esteem.

Several authors emphasize the need to consider how vulnerability mediates teachers' responses to reform (Kelchtermans, 1996; van Veen et al., 2005). The relationship between teachers' beliefs and social contexts within a climate of evaluation reform is central to the broader study of how causal conditions shape the decisions teachers make about their practices. Because of the inherent ethical and professional vulnerability of teaching, the effect of vulnerability on how teachers respond to policy, and the increased vulnerability created by value-added evaluation, particular attention must be paid to the role of vulnerability in teachers' adaptations to value-added evaluation policy.

Relational work. Elementary teachers especially claim not only to have affection for students but also, in some cases, even to love them (Nias, 1989; Hargreaves, 1994). Lortie (1975) called feelings of love and affection for students "psychic rewards of teaching" which are gained through classroom events and relationships with students (p. 187). Caring relationships with students are a source of professional satisfaction for teachers (Lortie, 1975). However, teachers do not experience gratifying, caring relationships without a commitment to relational work. Noddings (1992) suggested that caring requires teachers to elicit information from students and to listen to how students are feeling which in turn allows teachers to evaluate their purposes, help them to engage in self-evaluation, and help them grow as participants in caring relations. Caring can be seen both as an approach and as an emotion, and as such, it requires not only love but labor (Acker, 1995). The social and emotional aspects of teachers' relational work makes them vulnerable to reforms that reduce the time they have to develop caring relationships

with their students (Day, 2002). Hargreaves (1998) advised that when structures create a conflict between professional obligations and their ability to sustain their emotional commitments to their students, feelings of increased guilt and burnout can cause an exodus of teachers from the profession.

Identity “directly influences one’s emotional reactions to reform” (Reio Jr., 2005, p. 992). Day (2002) suggested that teachers’ sense of professional and personal identity is a key variable in their motivation and commitment to reforms, but more research is required to analyze how teachers’ identities are affected in the context of reforms. Hargreaves (2000) suggested that education policy, school administrators and most of the educational research community pay little or no attention to the emotions of teachers. Reio Jr. (2005) clarifies “The reform effort must take into account that teachers have natural emotional reactions to change that have both positive and negative influences on the construction of their professional and personal identity” (Reio Jr., 2005, p. 992).

Literature Review of Policy

The Expansion of Free Market Ideologies into the Public Sphere

Globalization is integral to the growth of capitalism (Hill, 2007). In *The Communist Manifesto*, Marx described globalization as a critical trait that enables capitalism to counter the tendency of profit rates to fall as wages rise, which results in a crisis of overproduction (Allman & Wallis, 1995). Since World War II, advancements in technology and the perceived failures of Keynesian economics have given rise to growing competition for global markets and investments (Lipman, 2007). The latest structural

reorganization of capitalism has led to changes in the function of the state, from the delivery of public services, to the promotion of private sector production and the management of a privatized public sector (Foley, 1994). The survival of capitalism depends upon growth; the shifts in the state's function are driven by capitalism's efforts to reproduce itself through new investment opportunities and sources for profit around the globe especially in industrially underdeveloped countries (Giroux & Giroux, 2006; Lipman, 2007). Many western governments and agencies have pursued a wide range of "deregulations, privatizations and abdications of responsibility to the market and private philanthropy," (Steinmetz, 2003 p. 337) including selling state-owned enterprises, goods, and services (such as public education) to private investors for increased efficiency and growth (Gibson & Ross, 2007). To further deregulate private industry and defund public services, governments and business leaders have aligned to endorse "a complex of values, ideologies, and practices that affect the economic, political, and cultural aspects of society," known as neoliberalism (Gibson & Ross, 2007, p. 1; Hursh, 2007). Neoliberal ideologies and policies provide the core ideological assumptions that underlie the current form of capitalist restructuring and globalization (Mayo, 2003). In her book *Neo-Liberal Ideology: History, concepts and policies*, Turner (2008) defined the key principles of neoliberalism regarding the market, the state, and the individual.

The market. Turner (2008) described the market within neoliberal thought as "an indispensable mechanism for efficiently allocating resources and safeguarding individual freedom" and "unfettered markets produce a natural order in society from the voluntary exchange of goods and services, promoting productive efficiency, social prosperity and

freedom” (p. 4). This is consistent with Adam Smith’s (1776) contention that “if any branch of trade, or any division of labour, be advantageous to the public, the freer and more general the competition, it will always be the more so” (Chapter II). Chubb and Moe (1990) also insisted that free markets are more efficient for allocating goods and services, more equitable because they are more responsive to the needs and wants of their clients than forms of monopolies in the public sector, and more democratic because they maximize freedom of individuals to choose interventions.

The state. Turner (2008) defined neoliberal ideologies regarding the state: “The liberal state should be strong but minimal: it should embody political authority but at the same time be constitutionally limited,” and its restricted role “is to secure social cohesion and stability through the preservation of individual liberties” (p. 5). Reflecting on the expansion of the government in the United States in the post-World War II era, economist Milton Friedman (1980) cautioned,

... the experience of recent years ... raises a doubt whether private ingenuity can continue to overcome the deadening effects of government control if we continue to grant ever more power to government ... Sooner or later - and perhaps sooner than many of us expect - an ever bigger government would destroy both the prosperity that we owe to the free market and human freedom proclaimed so eloquently in the Declaration of Independence. (Friedman & Friedman).

Proponents of privatization assert that reducing the role of the state and expanding the role of the free market provides a more efficient means to manage society than policies that seek to lessen income inequality (Tienken, 2013). Privatization policies are designed to liberate private enterprise from restrictions imposed by the state to reach an ideal of total freedom of movement for capital, goods, and services (Gibson & Ross, 2007, p. 3).

The individual. Lastly, Turner (2008) defined neoliberal ideology regarding the individual: “A system of full private ownership forms an indispensable part of a neo-liberal social order, reinforcing the irreplaceable value of the individual against the collective ... decentralizing decision-making and for placing it at the level of the individual” (p. 5). This echoed Adam Smith’s (1776) conception of the individual:

Every individual ... intends only his own gain; and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention. By pursuing his own interest, he frequently promotes that of the society more effectually than when he really intends to promote it. (Book IV, Chapter II).

Chubb and Moe (1990) asserted that the market is a tool to create choices that will lead to autonomy and that markets enable all participants to make decisions for themselves.

Because governments are less responsible for the welfare of the individual, the individual becomes responsible for him or herself. (Hursh, 2007). Free market, capitalist governments have shifted social responsibility from the community to the individual thereby minimizing the role of “the public good” in society (Gibson & Ross, 2007, p. 3) and redefining the relationship between the individual and society (Hursh, 2007).

Deregulation, privatizing public services, outsourcing, opposition to collective bargaining, and the elimination of tenure to lower labor costs are hallmarks of neoliberalism (Bourdieu, 1999). Governments promote neoliberal policies through a discourse of individual accountability, efficiency, and choice (Hursh, 2007; Lipman, 2007). These policies shift public funds and decision-making out of the public sphere and into private markets in the name of national interests. The broad acceptance of neoliberal ideologies has established “the superiority of free markets over public ownership” as the

new “conventional wisdom” from both ends of the political spectrum (Aronowitz, 2003, p. 21).

The Genealogy of Value-Added Evaluation in Education Policy

Mixed models have long been used in agriculture to analyze data to determine the effect of different variables on crop yields. It was William “Bill” Sanders, a statistician with the University of Tennessee, who applied these mathematical tools to isolate the impact of instruction from other variables on student learning (Ewing, 2011). *The Tennessean* (2017) reported that Sanders created and led a consulting group for the Institute of Agricultural Research for The University of Tennessee system, served as an adjunct professor of statistics for the College of Business, and eventually became the director of the Value-Added Research and Assessment Center at The University of Tennessee. While in Tennessee, Sanders and his colleagues investigated how covariate and mixed models used in agriculture could be adapted for new applications in educational assessment. In 1982, the Tennessee Department of Education commissioned his research on student and teacher data in Knox County and the resulting methodology, TVAAS (Tennessee Value-Added Assessment System) linked student academic outcomes to educational evaluation for the first time (Aldrich, 2017). While teachers in Tennessee began using value-added data in 1997, the state did not embed Sanders’ TVAAS within teacher evaluations until federal legislation catalyzed a need for new accountability measures for teachers and schools in 2009.

Applications of value-added models (VAM) in K-12 teacher evaluations rapidly spread across the United States, reifying a paradigm shift in public education reform that began decades before Sanders developed his TVAAS. In 1983, the United States National Commission on Excellence in Education submitted its report, *A Nation at Risk*, to the United States Department of Education (1983). In the report, the commission condemned the quality of American schools for their decline against historical performance measures and international standards. Provocative language compared the mediocrity of the public school system to an “act of war” (National Commission on Excellence in Education, 1983, p. 5). The American public school system appeared to have failed to improve over time and even worse, it was falling further behind other industrialized nations. The report’s message was clear: America’s poor educational system was causing the nation to lose its competitive advantage.

Over the next two decades, the perceived crisis in public education gathered momentum in the public and political discourse. Berliner and Biddle (1996) refuted many of the claims of public education’s critics with a re-examination of longitudinal studies of student performance measures. In fact, they found large increases in the average performance on the National Assessment of Educational Progress (NAEP) between the 1970s and 1990s for minority students, and stable or growing averages for all students (Berliner & Biddle, 1996). These counter-claims did not change the opinion of lawmakers or the public and Congress reauthorized the Elementary and Secondary Education Act (ESEA) of 1965 with the No Child Left Behind Act (NCLB) in 2001. The stated purpose of this watershed legislation was “to close the achievement gap with

accountability, flexibility and choice, so that no child is left behind” (U.S. Congress, 2002). Among other accountability mandates for states and school districts, NCLB required schools to show annual gains in students’ standardized test scores or annual yearly progress toward a specified academic proficiency standard (M. S. Tucker, 2014; U.S. Congress, 2002). School “report cards” based on standardized test scores were publicly released in an effort to increase accountability; failing schools could be closed, managed, or re-staffed by state-level agencies. Schools were explicitly accountable for lagging test scores and any achievement disparities between students of different backgrounds. This re-authorization marked a stark departure from the original ESEA that was passed over thirty-five years earlier in response to a similar crisis in public education; notably, differences in the political and social climate at the turn of the 20th century produced a much different legislative response to the perceived crisis in education.

Congress passed the original ESEA in 1965 to authorize federal funds (Title I) that explicitly addressed the social and economic inequalities found to negatively impact educational outcomes. The year before, the Civil Rights Act of 1964 commissioned The Equality of Educational Opportunity study. Widely referred to as the Coleman Report, the study offered a sociological review of disparities in student achievement based on a sample of over 600,000 students, teachers, and principals from across the U.S. (Coleman et al., 1966). The report provided data that clearly documented alarming disparities between average achievement levels of black and white students, still referred to today as the “achievement gap” (Bartz, 2016). The study also found that family background was

more closely correlated to a child's achievement than school resources, including school expenditures, class size, and teacher qualifications (Coleman et al., 1966). The authors' conclusion that students' socioeconomic backgrounds had a greater effect on achievement than school funding led to changed public opinion and public policy regarding the role of schools as the equalizer for children from disadvantaged socioeconomic environments (Bartz, 2016). Schools would receive significant financial support from the federal government through Title I grants based on the economic conditions of the students they served as a means to mediate the negative effects of poverty on student achievement.

However, by the time *A Nation at Risk* was released nearly two decades later, these federal funding efforts were deemed a failure by both political conservatives and social progressives in Congress. Therefore, the political response to the perceived crisis in education in the 1980s looked significantly different from the ESEA of 1965. Following the War on Poverty of the 1960s, accusations of impotence of the public sector had largely replaced the civic optimism of the mid 20th century. Education reformers, frustrated with the failure of schools to overcome the effects of poverty, looked to the corporate model for policy design; accountability in education came to mirror industrial values of productivity (Nichols & Berliner, 2007). Policymakers shifted the focus of reforms from addressing the socioeconomic inequalities among students to monitoring individual schools' effectiveness, leading to what Reeves (2005) called results-driven accountability systems.

In 2001, NCLB set in motion a monitoring system of standardized testing to hold individual schools accountable using the corporate model. Lipman (2007) suggested that the legislation looked to students' standardized test scores (outputs) "to serve as a surrogate for productivity" (2007, p. 36) by quantifying the quality of schools (production) and determining their value efficiently. Schools were rewarded or penalized based upon their calculated value. Differences between students (inputs), including social and economic conditions of students and school communities, were still considered in the NCLB-era evaluation models. However, reforms relied more upon improving the quality of schools to minimize differences among standardized test scores (the outputs) and less on addressing the inequalities of social and economic conditions among students (the inputs).

Value-Added Models in Tennessee Educator Evaluation Policy

Under NCLB, all schools were required to reach rates of 100% proficiency by 2014. As the deadline approached, states continued to struggle with persistent achievement gaps and stubborn proficiency rates, and the federal government responded with further market mechanisms to improve student test scores. President Barak Obama announced the Race to the Top (RTTT) federal grants in November of 2009. The \$4.35 billion fund would reward states that could show plans for "creating conditions of education innovation and reform" (U.S. Department of Education, 2009, p. 2). The RTTT grant was an opportunity for states to win millions of dollars in the midst of the Great Recession following the financial crisis of 2007-2008. Leveraging the financial conditions of cash-strapped states that were struggling to balance shrinking operating

budgets and fund schools at a time of lower tax revenues, the Obama administration attracted all but ten states to compete for much-needed assistance through market-based reforms (M. S. Tucker, 2014). Hoping to be selected for the multimillion-dollar award, thirty-nine states plus the District of Columbia initially applied for the grant, many with new teacher evaluations to motivate their educators to run harder and faster to the finish line (McNeil, 2010).

While NCLB introduced the industrial production principles of efficiency and competition to reforms, President Obama's RTTT program further integrated free market values and education reform policy. The language of the program's title elicits an image of a competitive race course and finish line at the "top" or summit of education reform. A "race" implied that the superior states would arrive at the "top" to win the competition, the less capable would fall behind and lose.

RTTT proposals were evaluated on a set of six selection criteria with 19 subcategories (see Table 2); the category, "Great Teachers and Leaders" weighed heavily in evaluating the proposals, accounting for the largest share with 138 points, or 28% of the total 485 possible points (2009, p. 3). Specifically, the subcategory, "Improving teacher and principal effectiveness based on performance," was ranked second among the nineteen criteria listed, with 58 possible points (2009, p. 3). The program explicitly valued and rewarded a teacher's "performance," on standardized tests, crystallizing the free market principles of individual responsibility and standardized meritocracy in education reform.

Table 2. Race to the Top proposal selection criteria

Race to the Top Selection Criteria
A. State Success Factors (125 points)
A.1. Articulating State's education reform agenda and LEA's participation in it (65 points)
A.2. Building strong statewide capacity to implement, scale up, and sustain proposed plans (30 points)
A.3. Demonstrating significant progress in raising achievement and closing gaps (30 points)
B. Standards and Assessments (70 points)
B.1. Developing and adopting common standards (40 points)
B.2. Developing and implementing common, high-quality assessments (20 points)
B.3. Supporting the transition to enhanced standards and high-quality assessments (20 points)
C. Data Systems to Support Instruction (47 points)
C.1. Fully implementing a statewide longitudinal data system (24 points)
C.2. Accessing and using State data (5 points)
C.3. Using data to improve instruction (18 points)
D. Great Teachers and Leaders (138 points)
D.1. Providing high-quality pathways for aspiring teachers and principals (21 points)
D.2. Improving teacher and principal effectiveness based on performance (58 points)
D.3. Ensuring equitable distribution of effective teachers and principals (25 points)
D.4. Improving the effectiveness of teacher and principal preparation programs (14 points)
D.5. Providing effective support to teachers and principals (20 points)
E. Turning Around the Lowest-Achieving Schools (50 points)
E.1. Intervening in the lowest-achieving schools and LEAs (10 points)
E.2. Turning around the lowest-achieving schools (40 points)
F. General Selection Criteria
F.1. Making education funding a priority (10 points)
F.2. Ensuring successful conditions for high-performing charters and other innovative schools (40 points)
F.3. Demonstrating other significant reform conditions (5 points)

(U.S. Department of Education, 2009, p. 9)

Following the introduction of RTTT in 2009, states with proficiency rates below 100% sought other means to avoid the impending penalties mandated by NCLB. The federal government offered NCLB waivers to states that passed legislation consistent

with the criteria of RTTT; many developed evaluations that held teachers and principals accountable for student scores on standardized tests. Tennessee was among the states that answered the call to compete for federal funding with the Tennessee First to the Top Act of 2010 (FTTT). In January of 2010, TN General Assembly passed a bill requiring school districts to use Tennessee's Value-Added Assessment System (TVAAS) and other student outcome measures as a significant component of principal and teacher evaluations starting in the 2011-12 school year (Piro, Wiemers, & Shutt, 2011). The state's efforts proved successful, earning Tennessee second place in the first round (Phase 1) of the RTTT competition and an award in excess of \$500 million (U.S. Department of Education, 2010).

The FTTT mandated that Tennessee school districts would incorporate student achievement scores in teacher and principal evaluations. Observation scores by state-certified evaluators using the TEAMTN rubric and other qualitative measures accounted for 50% of evaluation criteria using a five-point scale (1-5). The remaining half of evaluation scores was determined by two categories of student achievement. Thirty-five percent was based on student growth, which was determined by an individual teacher's TVAAS data for his or her students or a comparable measure (school-wide data were used for teachers without tested students). The remaining 15% came from other measures of student achievement adopted by the State Board of Education and mutually agreed upon by the educator and evaluator (Tennessee Department of Education, 2012). For example, a weighted calculation for a composite score for a teacher with tested students could include an observation score of 4, which would contribute half of his or her

composite score, a TVAAS score of 3, which would contribute 35% of the total score, and a score of 5 for the 15% achievement category. Weighting the three scores would result in a composite evaluation score of 3.8, or $(4 \times .5) + (3 \times .35) + (5 \times .15)$ (see Table 3 for example). The law required that districts use the composite score from annual evaluations for promotion, retention, tenure, and compensation decisions of all teachers and principals.

Table 3. Example of calculation of teacher effectiveness score in Tennessee

Teacher Effectiveness Score (1-5)		Scoring Example
Observation scores of effectiveness in teaching domains using TEAMTN Educator Rubric	Planning, Instruction, Environment, and Professionalism: 50%	Un-weighted Score = 4 Weighted Score = $4 \times 50\% = 2$
Student scores on mandated, standardized tests (TCAPs, TNReady)	Annual Growth: 35%	Un-weighted Score = 3 Weighted Score = $3 \times 35\% = 1.05$
	Achievement: 15%	Un-weighted Score = 5 Weighted Score = $5 \times 15\% = .75$
Overall Effectiveness Score	Total: 100%	Total weighted score = $2 + 1.05 + .75 = 3.80$

The teacher assessment system within Tennessee's FTTT legislation reflected the free market policy environment in which it was designed. The successful application of value-added models (VAM) in Tennessee's teacher evaluations rested upon the belief that leveraging competitive market forces and standardized meritocracies was the most efficient reform strategy for improving teacher quality and ultimately student outcomes. Policymakers drew from widely-cited research supporting teacher effectiveness as the

most influential school level factor in student achievement (Hanushek, 2014; Hattie, 2012; McCaffrey et al., 2004b; Rivkin et al., 2005; Stronge, 2007) even outweighing a student's background. Rivkin et al. (2005) reasoned that studies reveal "large differences among teachers in their impacts on achievement and show that high quality instruction throughout primary school could substantially offset disadvantages associated with low socioeconomic background" (p. 419). This influential body of research led policy makers to design mandates intended to directly impact teachers' practices and elicit more "effective" behaviors. Hursh (2007) claimed that value-added evaluation systems like FTTT assumed that a teacher is a "competitive, instrumentally rational individual" and employed competitive conditions and high stakes incentives and consequences to motivate teachers to engage in more "effective" teaching behaviors (p.16). The underlying market principles embedded in applications of value-added models in teacher evaluations looked to the individual efforts of teachers to overcome the effects of social and economic inequality that influenced academic gains and caused achievement gaps. In response to persistent inequality and stagnant proficiency rates in educational outcomes since the adoption of ESEA reforms, the RTTT program and the value-added evaluation systems it spawned across the country were designed to leverage market principles of competition and meritocracy to motivate individual teachers and principals to address gaps and shortfalls in achievement. NCLB, RTTT, and FTTT have marked a departure from the ESEA legislation of 1965 which targeted the social and economic causes of educational inequalities through systemic changes in educational funding and resource reallocation (e.g. Title I, II, and III funding to support schools that serve economically

disadvantaged students, lack highly qualified teachers, and provide English language acquisition instruction, respectively.)

Free Market Ideologies in Teacher Evaluation Policy

Helsby (1999) contended that since World War II, governments have attempted to reposition the strong liberal-humanist view of schooling, characterized by a belief in the intrinsic, non-instrumental value of education, towards a more functional view that embraces competency based, results-driven teaching. As a result of the larger economic and political shift toward free market ideologies in the public sector, high stakes testing systems are now used in public schools to ensure that all students meet prescribed levels of achievement on state standardized tests in the U.S. (Day, 2002). Hallinger et al. (2014) suggested that as the global context has shifted toward a preference for private management, conceptions of accountability in education have changed as well.

Accountability is not simply a set of policies; it is an institutional logic of a rationalized ideal that models how schools should operate (Hallett, 2010). Cochran-Smith et al. (2018) suggested that accountability is now “a major strategy” for improving K-12 education, higher education, and the public sector through a discourse of “heightened auditing, monitoring and surveillance” (p. 9). Tuytens and Devos (2014) claimed that in many countries, a “new public management” (p. 156) has introduced techniques from the private sector to bring an increased focus on accountability to education. This management approach is not value-free because it has an implicit and explicit set of values based on an economic rationale and conception of people as human resources who

need to be managed to maximize profits (Ruggunan & Spiller, 2014). In an effort to increase their economic competitiveness, governments use policy mandates to relocate accountability for student outcomes from schools to individual teachers (Day, 2002; Hallinger et al., 2014). With a standardized “delivery system,” Berlak (2000) suggested that authorities can identify “teachers, schools and local districts that fail to produce, and institute marketplace remedies, privatization, vouchers, charter schools and other policies that encourage schools to compete for students and resources” (p. 190).

While Democrats in Congress historically supported teachers and labor unions, by the end of the 20th century, their resolve to address achievement gaps with federal funding had weakened under pressure from increased global competition, and schools were failing to deliver a workforce that could meet the needs of the national economy (Tucker, 2014). Both conservatives and social progressives, frustrated with the failure of past investments in public education to close the achievement gap, widely supported Obama’s RTTT that required new levels of individual accountability for teachers and other reform policies based upon market competition (charter schools) and free-market systems (merit-based pay) (Aronowitz, 2003). For Democrats, emphasizing standards was a way to promote equality, while Republicans viewed accountability as a means to improve transparency and efficiency (Mehta, 2008 cited in Hallett, 2010).

RTTT proposals and NCLB waivers incorporated sophisticated statistical models designed to isolate, measure, and standardize the value individual teachers contributed to his or her students’ academic gains. NCLB and RTTT are both grounded in the global policy climate of the early 21st century where free market competition and meritocracies

were the framing principles reforming public education and other publicly managed institutions (Giroux & Giroux, 2006). Inter-dependent ideologies of free markets and standardized meritocracies played an important role in processes of education reform and aligning schools with methods, cultures, and ethical systems of the private sector (Ball, 2003).

Consistent with these free market principles, William Sanders (1994), of the University of Tennessee, developed the Tennessee Value-Added Assessment System (TVAAS) model to measure “the product of the educational experience rather than the process by which it was to be achieved” (p. 300). In education, value-added was the standardized, quantitative contribution of a specific teacher to a specific student’s academic gains. Students’ past and current academic outcomes, measured by standardized tests, were compared to determine a growth rate. This rate was analyzed against other students’ growth trajectories over the same time period to determine how a teacher’s contribution ranked among other comparable teachers across the state to determine his or her value-added score. Supervisors were required to include these scores for hiring, firing, and other personnel decisions, including salary reduction, loss of tenure, or denial of merit pay.

According to Sanders and Horn (1994), “TVAAS was developed on the premise that society has a right to expect that schools will provide students with the opportunity for academic gain regardless of the level at which the students enter the educational venue” (William L Sanders & Horn, 1994, p. 301). This quantitative analysis of productivity applied to teachers and students shaped the narrative of education reform to

focus on improving individual teacher performance, largely ignoring systemic social and economic inequality. Education reform since RTTT shifted the values and discourse around the factors that influenced educational outcomes to focus nearly exclusively on teacher effectiveness. Developed within an environment of free market hegemony, teacher evaluation policies based on value-added scores looked to principles of market competition and individual responsibility to improve teacher quality. These policies were designed to motivate teachers to improve the quality of their practice using a market system of rewards and consequences.

Social and economic intervention policies of the 1960s were designed to address the effects of poverty and socioeconomic background on student achievement. In fact, over the past fifty years, education reforms in the United States focused consistently on improving academic achievement for all students; education policy design, however, has evolved over time to reflect the shifting sociopolitical climate surrounding public education. When federally funded programs of ESEA appeared inefficient and ineffective to achieve equity in educational outcomes, policymakers turned to free market principles and the industrial production model for reforms, resulting in NCLB funding that required states to monitor the performance of schools through standardized testing systems. However, when school level accountability proved inadequate to achieve the student proficiency goal of 100%, education reformers returned to market principles of standardized meritocracy, competition and individual responsibility and accountability. RTTT funding and NCLB waivers reified these ideologies as impetuses for evaluation policy designed to control for all extraneous factors and to place the responsibility of

student achievement squarely on the shoulders of individual teachers. Half a century after the Coleman Report and the creation of Title I funding in 1965, paradigms of meritocracy and free-market efficiencies have come to dominate education reform, resulting in evaluation policies designed to directly impact teacher behavior using a competitive accountability system of high-stakes consequences.

Empirical Research on Value-Added Methods

Potential Benefits and Limitations of Value-Added Methods

Educational value-added assessment systems use multivariate, longitudinal modeling to assess the effectiveness of districts, schools, and teachers and to provide dynamic projections of student performance and needs (Hagstrom, 2015). Annual yearly progress (AYP) models of school performance reflect the percentage of students who have scored at or above a particular proficiency score. Although these measures are simple and transparent, “a percent proficient indicator is, in many ways, one of the weakest indicators of performance” (Choi, Goldschmidt, & Yamashiro, 2005, p. 9). Braun (2005) suggested that value-added models (VAM) may be the only way to conduct fair teacher evaluations because teachers and schools are evaluated based on their contributions, not an absolute standard of proficiency. Amrein-Beardsley (2008) suggested that value-added models can provide a “more defensible” method to measure a teacher’s effectiveness “than by simply relying on a traditional ‘snapshot’ measure ... capturing the level at which students exited the classroom independent of their level when entering” (p. 65).

Proponents of statistical models that measure a teacher's value also claim that its use in teacher evaluations can improve instruction and student outcomes. The goal of VAM is to allow educators and policymakers to make apples-to-apples comparisons among teachers in terms of how much content their students learn each year regardless of the students' characteristics (RAND Corporation, 2002). A report by the Education Trust suggested that value-added data provide administrators with information to improve teacher effectiveness (Jerald, 2009). The Tennessee Department of Education (2012) credited VAM in teacher evaluations with the fastest growing achievement scores the state has seen in any previously measured year. Sanders and Horn (1998) insisted that without value-added information "educational improvement efforts cannot address the real factors that have been proven to have the greatest effect on student learning" (p. 256).

However, the National Research Council questioned the validity of such claims and advised that assessments "that mimic the structure of large-scale, high-stakes, summative tests, which lightly sample broad domains of content taught over an extended period of time, are unlikely to provide the kind of fine-grained, diagnostic information that teachers need to guide their day-to-day instructional decisions" (2009, p. 10). Since the implementation of TVAAS and other value-added models, numerous studies have investigated the validity and reliability of these statistical models in teacher evaluation policy with varied conclusions and policy recommendations. Proponents of Tennessee Value-Added Assessment System (TVAAS) and other value-added statistical models claim their methods successfully control for students' characteristics and socioeconomic

backgrounds. Sanders, the architect of the TVAAS, acknowledged that to assess teacher influence on student learning, schools and school districts must control for the many confounding variables that affect students' academic gains, yet schools and districts do not have access to enough data required to do so (William L Sanders & Horn, 1994).

With the TVAAS model, Sanders and Horn (1994) asserted that "influences can be filtered without having to have direct measures of all the concomitant variables" (p. 305). Rivkin et al. (2005) noted that "repeated performance observations for individual students and multiple cohorts provide a means of controlling explicitly for student heterogeneity and the nonrandom matching of students, teachers, and schools through the use of fixed effects models" (p. 418). Instead of attempting to adjust for all factors not related to teacher influences, the authors explained that the TVAAS model focuses entirely on measures of academic gain so that each child serves as his or her own "control" (William L Sanders & Horn, 1994). Choi, Goldschmidt, and Yamashiro (2005) suggested that a student's initial achievement score captures many of the socioeconomic effects that the model is attempting to measure. Therefore, the model already captures the potential effects of a students' socioeconomic status.

However, some critics of VAM in teacher evaluations draw attention to the impact of school-level factors that compromise internal validity, reasoning that VAM cannot control for the non-random assignment of students (Darling-Hammond, Amrein-Beardsley, Haertel, & Rothstein, 2012; McCaffrey et al., 2004b; Piro et al., 2011). McCaffrey et al. (2004b) insisted that student characteristics cannot be isolated from teacher effects on student performance, especially when comparing groups of students

from “distinctly different populations” (p. 1). Baker et al. (2013) claimed that when students are not randomly assigned to teachers, inferential statistics are not applicable. Braun (2005) also explained that using value-added scores as a measure of a teacher’s contribution to student learning is “equivalent to making a causal interpretation of a statistical estimate” and counseled that unless students are randomly assigned to teachers and schools, “causal interpretations can be misleading” (p. 3). McCaffrey et al. (2004a) suggested that research investigating the correlation of VAM measures of teacher effects with alternative indicators of teacher effectiveness, including principal observations and other qualities of effective teachers, is necessary for criterion-based evidence of the validity of VAM.

Darling-Hammond et al. (2011, 2012) also suggested that value-added ratings are highly dependent on the students they teach. Research suggests that low-income students begin their schooling less prepared to succeed in a standards-based education system than their more affluent peers (Risley & Hart, 1995). These students achieve lower scale scores and lower rates of proficiency than their less-impooverished peers (Tienken, 2013). In her review of student growth scores in Ohio, Pennsylvania, and Arizona, Amrein-Beardsley (2018) found that a disproportionate number of high-poverty schools score low value-added scores compared to schools with higher average incomes. Welsch and Zimmer (2010, p. 46) observed that poverty is positively related to children’s reading performance. Risley and colleagues (1995) also offered evidence that poverty is associated with lower test scores because students from low-income homes begin school less prepared to master state standards. In their longitudinal study, Risley and Hart (1995)

found that the cumulative differences between the vocabularies of children at age 3 in families on welfare and those from working-class families were predictive of the children's reading performance at ages 9-10. In another longitudinal study, Raver et al. (2013) found that "chronic exposure to poverty and the strains of financial hardship were each uniquely predictive of young children's performance on measures of executive functioning" (p. 292). In a national study of 400 Chapter 1 schools, researchers found that higher poverty, greater application of grade-retention policies, and more student disciplinary actions were related to lower student achievement (Puma et al., 1997). In its letter to the U.S. Department of Education, the Board of Testing and Assessment (2009) also recommended against using value-added evaluation because of variables that cannot be isolated or controlled: "VAM estimates of teacher effectiveness should not be used as the sole or primary basis for making operational decisions because the extent to which the measures reflect the contribution of teachers themselves, rather than other factors, is not understood" (p. 10). Darling-Hammond et al. (2011) also cautioned that even when controlling for socioeconomic and prior achievement variables, teachers of English language learners and students with special needs show lower growth scores than when they are teaching a different population.

The accuracy of value-added models is also in question. Schochet and Chiang (2013) reported that the number of years of data used in value-added models can significantly affect the error rates for individual teacher effectiveness scores. For example, error rates can cause one in four teachers to be misclassified with three years of student test data, and one in three will be misclassified when using a model that considers

only a single year (Schochet & Chiang, 2013). In a study of more than 12,700 teachers, Corcoran (2015) found large margins of error in both math (34 percentile points) and English language arts (44 percentile points) value-added scores, making it difficult to use these scores to measure teacher effectiveness with confidence. The American Educational Research Association (AERA) also identified similar limitations of using VAM to evaluate educators. In an official statement, the AERA Council (2015) contended that the validity of standardized tests varies because not all “fully capture the target constructs” or measure student achievement with reliable precision (p. 449). Therefore, the quality of data for calculating VAM results must be constantly monitored when used for educator evaluation.

Model errors aside, testing decisions made by districts and schools can also impact individual teacher value-added scores unintentionally. In the executive summary of their study of value-added evaluation in the Los Angeles Unified School District, Briggs and Domingue (2011) found that teacher effectiveness ratings were “quite sensitive to choices” among the value-added models used to calculate scores (p. 3). Papay (2011) advised that variability of the achievement tests used and the timing of tests during the school year can contribute to instability of value-added estimates used in teacher evaluations. Another report compared the scores from tests using open-ended questions to those using selected response questions and found low correlations between the two, suggesting that the format of test questions can significantly impact student achievement measures (Bill and Melinda Gates Foundation, 2010). The reliability and

validity of these measures affects the accuracy of the value-added models that calculate teacher effectiveness.

Error, bias, and uncontrolled factors in student testing negatively affect the accuracy of student achievement measures; value-added models with incomplete inputs compromise the validity of teacher effectiveness measures. Other critics challenge the exclusive use of student growth scores to measure a teacher's value-added, or contributions to student learning. Grissom, Loeb, and Doss (2015) suggested that teacher evaluations that give too much weight to student growth scores are "likely to miss important contributions that 'low-value-added' teachers make to their schools (p. 48). Jackson (2012) found only a weak correlation between teachers' value-added scores to students' non-cognitive outcomes, suggesting that value-added scores can fail to capture key aspects of teachers' impact on students. There is a "difference between process of learning and process of demonstrating mastery of concepts on a single test administered on a single spring day" (Tienken, 2013, p. 311).

While the weaknesses of VAM models do not render them useless, Gabriel and Lester (2013) counseled against using value-added scores as a significant measure of teacher effectiveness, especially when job security and compensation decisions are at stake. Baker, Oluwole, and Green (2013) insisted that VAM evaluation policies with practical and financial implications must rest upon "objective measures of student achievement growth" (p. 5). The National Research Council (2009) suggested "VAM estimates of teacher effectiveness that are based on data for a single class of students should not be used to make operational decisions because such estimates are far too

unstable to be considered fair or reliable” (p. 10). Similarly, Gabriel and Lester (2013) asserted that when VAM are used in high stakes environments to make decisions about merit pay, promotion, retention, and termination, the tolerance for error and bias must be lower than in non-evaluative uses. Statistical errors in growth models, bias in test design, and uncontrolled testing conditions can produce unreliable calculations of student outcomes and of teacher value-added scores derived from them; these errors and biases can impact the professional future and financial security of teachers when teacher value-added scores are tied to retention, tenure, pay and hiring decisions.

The high stakes of value-added modeling in teacher and principal evaluations have sparked debates over the validity of those methods and applications in teacher evaluation. While the controversy over using value-added models in teacher evaluation is unlikely to end soon, critiquing its internal validity is a red herring for assessing the impact of these evaluation policies on what goes on in classrooms. If educational reformers intended to motivate teachers and influence their behaviors with value-added modeling in teacher evaluations, then those methods of accountability must be judged on how teachers have responded in practice. Spillane and Burch (2006) asserted that researchers are challenged with developing “an understanding of *what* people do, *how* they do it, and *why* they do it, while simultaneously attending to the institutional structures at various levels of the system that enable and constrain that activity” (p. 97). A review of research of teachers’ perceptions of value-added evaluation and teachers’ responses in practice is essential for understanding the impact of reforms in the practices of teachers. The next section discusses the extant literature on how teachers have

responded to value-added components in their professional evaluations and the impact of their responses.

Teachers' Responses to Value-Added Evaluation and Their Consequences

Hargreaves (2005) suggested that high-stakes testing and the standardization movement in education reform have narrowed the curriculum, destroyed creativity in classrooms, and encouraged “cynical and calculative strategies” for improving student test scores (p. 105). This is especially the case among teachers serving in high-poverty schools that are under pressure to meet minimum test scores and are compelled to narrow the taught curriculum and alter instruction and pedagogy to maximize test preparation (Au, 2011; Hursh, 2007; Nichols & Berliner, 2007). This pressure creates conflicts for teachers who must choose between meeting institutional expectations and requirements or the goals and requirements of best practices in teaching and learning (J. W. Meyer & Rowan, 1977). As curriculum is aligned with standardized assessments, “the focus is more and more only on those elements that can be easily measured on standardized tests. Knowledge *that* and occasionally low-level knowledge *how* are the primary foci” (emphasis added) (Apple & Jungck, 1990, p. 234). Teachers are compelled to adopt more teacher-centered, standardized pedagogies that are contradictory to constructivist, student-centered practices because they “deliver test-driven curriculum in an efficient manner” (Au, 2011, p. 31). Teachers increase the amount of time students perform test drills and practice for types of knowledge and test taking skills that the tests require (Au, 2011, p. 31). In a case study of an elementary school’s response to Texas’ accountability system, Booher-Jennings (2005) found that because special education students were not

included in school accountability ratings, teachers had their potentially low-scoring students tested to determine whether they qualified for special education.

High-stakes testing potentially diminishes the educational opportunities of all students because of the pressure it creates to narrow curricula to tested subjects and skills. However, its effects have been seen more frequently in low-achieving schools that are obligated to use test-preparation materials as texts, to narrowly focus on the tasks that are tested, to concentrate much class time on test-taking skills, and to reduce learning to passing the tests (Lipman, 2007). When teachers alter curricula and pedagogies to improve students' proficiency rates on standardized tests, they limit socioeconomically disadvantaged students' access to quality educational opportunities. Tienken (2013) described how low performing students have been offered fewer and more narrowly defined educational opportunities designed to improve test performance, which has led to a "cycle of educational austerity," perpetuated by teachers' adaptive practices (p. 305). Booher-Jennings (2005, p. 232) found that teachers in one Texas elementary school engaged in several practices as a form of "educational triage," a system of sorting students by their perceived potential to improve school test scores and allocating resources accordingly. In the study, teachers dedicated more resources for teaching students perceived to have the most potential for improving the schools' accountability rating, and fewer resources for students viewed as "hopeless cases" (Booher-Jennings, 2005, p. 233). While these adaptive practices were in teachers' best interests for improving their professional evaluation ratings, Tienken (2013) suggested that these adaptations also led to a Social Darwinism where only the best-adapted or fittest students

can succeed. With less access to high quality academic tracks, disadvantaged students spend more time in lower-level basic skills test preparation classes than their non-disadvantaged peers, disproportionately limiting the educational experience and opportunities of socially and economically disadvantaged students. Within the standardized meritocracy paradigm, the public education system is viewed as “the ultimate arbiter of innate intelligence and ability, as well as the benefactor of hard-work and merit” (Leyva, 2009, p. 365). When the ideology of standardized meritocracy inherent in high-stakes testing systems is accepted as a legitimate and efficient method to sort students and determine the quality of their education and future employment opportunities, educational opportunity for disadvantaged students can decline into a downward spiral.

Critics of VAM have suggested that the external validity of value-added models is compromised when they are applied in contexts with significant consequences for teachers. The competitive market principles and industrial production values inherent in high-stakes testing have created pressure on teachers to engage in less cooperative and even unethical behaviors (Nichols & Berliner, 2007). Campbell’s (1976) law suggested that under normal teaching conditions, standardized tests can provide valuable information for assessing academic outcomes. However, when the primary purpose of education becomes the production of test scores, a teacher evaluation policy can distort the measures to make them less valid or useful for improving instruction and creates several unintended consequences (Campbell, 1976; Collins, 2014; Nichols & Berliner, 2007). These outcomes include a growing disincentive to teach challenging student

subgroups and increased competition for non-disadvantaged students, less cooperation among teachers, and a rise in behaviors to maximize test scores at the expense of long term achievement gains (Collins, 2014; Corcoran, Jennings, & Beveridge, 2011; Johnson, 2015). Hursh (2007) claimed that free market competition in education can lead to the commodification of students (as inputs) because some pupils are judged to have more value over others in producing the best scores (as outputs). In her study of a value-added assessment system in the Southwest United States, Collins (2014) reported teachers' descriptions of how recruiting for class rosters with the best student profiles created a competitive atmosphere. In addition, some teachers self-reported "unprofessional and unethical behaviors" to improve their students' test scores (Collins, 2014, p. 18).

Education policy that creates a competition for test scores can undermine teachers' ethic of social responsibility to educate all students and can lead to behaviors that compromise teachers' relationships with their students.

Effective schools research suggested a strong correlation between collaboration among teachers and student achievement (DuFour, DuFour, & Eaker, 2008; Eaker & Keating, 2011; Lezotte & Snyder, 2010). However, cooperation is not measured or calculated in value-added assessments of teachers. Harris and his colleagues (2014) were unable to find more than a weak correlation between value-added scores and principals' ratings of teachers' characteristics related to team work (e.g. low value-added teachers did not necessarily receive high collaboration scores from their supervisors). The results suggested that rewarding an individual's value-added results over teamwork may attract "lone wolves" to the profession, (teachers who resist collaborative behaviors) influencing

the composition of personality types within the profession over time (Braun, 2015; Harris et al., 2014). Braun (2015) suggested that the de facto rating of teachers by their value-added evaluation scores may change the nature of their work by creating competition between teachers that undermines collegial relationships. Johnson (2015) proposed that diminished collaboration and interactions between colleagues can limit a school's ability to transform its human capacity into the social capital it needs for successful school improvement. He offered an organizational perspective to VAM, calling for more research on the effects of VAM on the organizational health of schools.

Teacher Perceptions

Since its introduction in 2009, Race to the Top (RTTT) funding and pressure on state legislatures to acquire waivers for the requirements of NCLB have catalyzed the widespread implementation of VAM evaluation policies for teacher accountability. Within the current educational climate, multiple studies have subsequently explored teachers' perceptions of VAM assessment systems. Overwhelmingly, teachers report distrust with the methods, and they report low morale as a result of perceived changes in their teaching to earn a successful professional rating (Collins, 2014; Feuer, 2012; Firestone, 2014; Goldhaber, DeArmond, & Deburgomaster, 2011; Jiang et al., 2015; Lee, 2011; Michalek, 2014; Tuytens & Devos, 2009). Educators, parents, and policy makers, among others, struggle with the lack of transparency of value-added models due to their high level of statistical sophistication (National Research Council, 2009). Jiang et al. (2015) suggested that many teachers have expressed confusion and misinformation about how much student growth contributes to overall evaluation scores and how VAM

controls for external influences including mobility and poverty. Similarly, Lee (2011) reported a mixed-method study in a large urban district suggesting that “what VAM measures is not aligned to what teachers see as the purpose of education” (p. 102).

The implementation of school policy is more than a rational or technical process; therefore, taking the emotions teachers experience into account is important (Zembylas & Barker, 2007). Schmidt and Datnow (2005) explained that because reforms are interpreted and re-interpreted by the teachers who implement them, teachers’ emotions are important factors in shaping their behaviors in response to reforms. In their study of teachers’ perceptions of a new evaluation policy, Tuytens and Devos (2009) found that teacher perception of policy is key to understanding the success or failure of policy. Van den Burg, Vandenberghe, and Slegers (1999) insisted that “it is not so much the objective characteristics of an innovation that determine the success of the innovation. Explanations and problems can stem from the meaning or significance that people assign to the new situation produced by the innovation” (p. 335). Teachers’ constructions of meaning shape how they respond to reforms, which are not necessarily in ways that policy makers intended (Fitz, 1994). Therefore, when analyzing the outcomes of a new policy, it is important to analyze its objective characteristics *and* to analyze the meaning teachers attach to policy and its characteristics (Geijsel, Slegers, van den Berg, & Kelchtermans, 2001).

This chapter was divided into three sections: A review of the conceptual framework, a literature review of policy, and a review of empirical research of value-added methods and teachers’ responses. The first section provided a literature review of

the sociocultural and ecological models used to frame the study. This section began with a discussion of relevant institutional theories for analyzing public policy implementation. Next, the section defined the scope of sensemaking theory, Bandura's conceptions of agency and the dimensions of teaching that impact sensemaking. This discussion provided a lens to explore the dynamics between institutional logics of value-added evaluation policy, teachers' perceptions of policy, and teachers' responses in practice. The second section provided a review of the extant literature of value-added evaluation policy including the socio-cultural sources of the ideologies embedded in value-added models. The section continued with a genealogy of value-added models in American teacher evaluation policy. A literature review of empirical research on value-added methods and on teachers' responses to value-added evaluation policy concluded the chapter. Chapter 3 will provide an explanation of the impetus and rationale for the study's design, a description of the methods, and a discussion of the criteria for the quality of the research.

CHAPTER III

METHODOLOGY

The purpose of this grounded theory study is to construct a theoretical model that explains the process teachers experience in choosing their practices in response to a state-mandated, value-added teacher evaluation policy.

The data collection for the study is guided by the following research question:

- What theoretical model can explain how teachers have adapted their practices in response to value-added evaluation policy?

From this central question, two empirical questions follow:

- What are the lived experiences of teachers who have been evaluated by the Tennessee Value-Added Assessment System?
- What practices do teachers choose in response to Tennessee Value-Added Assessment System evaluation policy?

This chapter provides an explanation of the impetus and rationale for the study's design, a description of the methods, and a discussion of the criteria for the quality of the research.

Impetus for Design

As detailed in Chapter 2, President Obama signed the American Recovery and Reinvestment Act (ARRA) into law in February of 2009. Among other stimulation measures, the legislation created a \$4.35 billion Race to the Top (RTTT) Fund to reward

states that implemented reforms that specifically used clearly established student growth measurements that could be used to evaluate teachers and principals. Proposals for the funding needed to include provisions that the state would use these evaluations for “compensating, promoting, and retaining teachers and principals,” and to determine “whether to grant tenure and/or full certification” (U.S. Department of Education, 2009). Tennessee was among the first states funded through RTTT and in the fall of 2011, public school districts across the state implemented the Tennessee Educator Acceleration Model (TEAM). The evaluation policy included the Tennessee Value-Added Assessment System (TVAAS) that tied a percentage of teacher and principal evaluation scores to their students’ annual growth in achievement scores on the Tennessee Comprehensive Assessment Program (TCAP), a set of standardized tests taken by all students in grades 3-12 each year. Public policy intended to achieve specific goals are “forms of intervention by public authorities” (Bemelmans-Videc, 2011, p. 4). Tennessee adopted TVAAS to meet RTTT “Reform Plan Criteria,” of “improving teacher and principal effectiveness through performance” (U.S. Department of Education, 2009). This value-added evaluation policy was designed to influence teachers’ and principals’ behaviors by linking their pay, job security, and career advancement to their students’ performance on standardized tests. Designed as an instrument to influence the behavior of teachers and principals, the stated purpose of the TEAM and TVAAS policies was to improve teacher quality.

As a teacher at a Title I school in Tennessee, one of the first states to receive Race to the Top funding, the researcher is intimately aware of the challenges that faced

administrators, teachers, and students in economically disadvantaged communities in an era of results-based accountability. The researcher and her colleagues are among thousands of public school teachers in the state to receive an annual score based, in part, on value-added methods. The researcher witnesses how public policy intended to improve students' education through improved teacher quality results in unintended consequences. These outcomes are compounded within economically disadvantaged schools, especially compared to the more affluent school districts in the same county. Nichols and Berliner (2007) suggested that Campbell's law explained many of the unethical and sometimes illegal behaviors of schools and teachers resulting from current accountability measures in public education. The researcher suggests that classroom teachers, specifically in economically disadvantaged communities, adopt additional subtle behaviors that are largely unexamined, and little understood to adapt to value-added evaluation policies. The researcher's experiences and observations as a classroom teacher since 2011 in Title I elementary schools in Tennessee are the impetus of the research design. This study explores the dynamics of how teachers adapt to value-added evaluation policy in elementary schools. Teachers' perceptions of this policy are unknown and the impact of the policy on teachers' agency and decisions in practice have gone unexamined. This study focuses on this gap by conducting research to address the question: What theoretical model can explain the process teachers experience when choosing their practices in response to value-added evaluation?

Thus, the study is designed to collect data through semi-structured interviews from teachers across a single public-school district within Tennessee. Using teachers'

responses and narratives, the report proposes a substantive process model to explain the adaptations of teachers using their own words and self-reported perceptions and behaviors. The model is intended to empower teachers responding to policy and shaping implementation through their daily choices in practice, to advise school and district level administrators implementing evaluation policy through their choices in institutional practices, and to enlighten decision makers selecting or designing teacher evaluation policy.

Rationale for Design

Integrity of Design

Marshall and Rossman (2011) suggested that a compelling research proposal presents a strong connection between the genre, overall strategy, research questions, design, and methods. The researcher's goal is to achieve epistemological integrity between the questions asked and the methods of data collection. The rationale for a constructivist grounded theory study that employs narrative inquiry and critical ethnographic data collection methods to generate a process model is outlined below.

Qualitative Rationale

The central research question addressed objectives specific to the local context of this inquiry; the researcher seeks to generate a grounded theoretical model that may explain how teachers adapt their practices in response to value-added evaluation. Describing and ascribing meaning to teachers' perceptions, conceptualizations, behaviors, and agency requires an interpretive approach to research (Creswell, 2012).

Semi-structured interviews provide a means to document teachers' accounts of their lived experiences, perceptions, and choices (Clarke, 2003; Norman K. Denzin & Lincoln, 2008). The absence of teacher voice in policy research warrants the use of qualitative methods to gather and represent participants' experiences in their own words.

Grounded Theory Design

This grounded theory study specifically collects data to describe the participants' experiences with value-added evaluation and the social contexts of those events. The design continued to emerge as the research was conducted, and the analysis and representation of findings are interpretive and include multiple perspectives (Creswell, 2012; Marshall & Rossman, 2011; Schwandt, 2007). Schwandt (2007) described, "Grounded theory methodology is a specific, highly developed, rigorous set of procedures for producing formal, substantive theory of social phenomena" (p. 110). Grounded theory owes its strengths to several critical components that are not necessarily included in other inductive methods (Bryant & Charmaz, 2007). Key elements include constant comparative analysis and theoretical saturation through theoretical sampling, and extensive memo writing (Bryant & Charmaz, 2007; Creswell, 2012; Glaser & Strauss, 1968). This study employs the following methods:

- **Constant comparative analysis.** According to Charmaz (2005), "Grounded theory is a comparative method in which the researcher compares data to data, data with categories, and category with category" (p. 517). The analysis compares interviews with each other and to open, selective, and theoretical categories throughout each cycle with consistent

memo writing (Corbin & Strauss, 2008; Creswell, 2012; Miles & Huberman, 1994).

- **Theoretical sampling.** Sampling for this study is directed by an analysis of the data collected beginning with the first interview and guided by the categories developed from ongoing memo writing and data analysis. The analysis of interviews continues reiteratively throughout the data collection phase to determine if a need for additional participants to reach theoretical saturation exists (Corbin & Strauss, 2008; Creswell, 2012; Maxwell, 2012).
- **Theoretical saturation.** Theoretical saturation is the point when no new data emerge and all concepts in the theory or model are well-developed from the sampling and analysis processes (Morse, 2004). Data collection continues to gather concepts that appear to be relevant to the theory until the model no longer needs to be modified to account for the findings. The sample size is determined when analysis reaches “theoretical saturation” of categories, instead of demographic representativeness (Corbin & Strauss, 2008; Creswell, 2012; Glaser & Strauss, 1968; Lempert, 2007; Maxwell, 2012; Miles & Huberman, 1994). Therefore, generalizing the findings from this sample differs from statistical generalization, and criteria for generalizing are theoretical rather than statistical (Bryant & Charmaz, 2007; Charmaz, 2014).

The analysis combines theoretical sampling and theoretical saturation of categories to develop a substantive (contextually bound) process model (Bryant & Charmaz, 2007). Charmaz (2005) argued that writing a grounded theory report should be “intuitive, inventive, and interpretive, not merely a reporting of act and facts, or...causes, conditions, categories, and consequences” (p. 529). At the same time, the report should produce an analytical product, not just a descriptive account. Therefore, this report provides an analytic framework to evaluate the findings and reflects the researcher’s emic interpretations as a participant of the critical reference group.

Constructivist Grounded Theory Approach

Grounded theorists including Glaser, Strauss, and Corbin have advanced positivist methods based upon objectivist assumptions, but by the end of the 20th century, a constructivist approach to grounded theory gained a following of researchers who rejected the modernist epistemology of earlier versions of the method (Charmaz, 2014). In fact, Corbin (2008) admitted that the influence of feminist and postmodern paradigms led to an evolution in her approach to analysis. Corbin (2008) agreed with the constructivist approach where “concepts and theories are *constructed* by researchers out of stories that are constructed by research participants who are trying to explain and make sense out of their experiences and/or lives, both to the researcher and themselves” (p. 10). Describing the role of constructivist research Schwandt (1998) explained, “constructivism means that human beings do not find or discover knowledge so much as construct or make it. We invent concepts, models, and schemes to make sense of

experience and, further, we continually test and modify these constructions in light of new experience” (p. 237).

A constructivist grounded theory study is effective for investigating participants’ ideas and beliefs expressed through their language or behaviors within the culture. With recent epistemological turns in traditional grounded theory methods, investigations of discourse, agency, structures, images, context, history, and current events are included to analyze complex social situations (Clarke, 2003). Constructivist methods look for patterns of social organization and worldviews and produce understandings of how culture-sharing groups function (Creswell, 2012). With the constructivist approach, grounded theory begins with the assumption that social realities are constructed and representations are interpretive through the researcher’s “position, privileges, perspective, and interactions” which must be disclosed and accounted for in the research analysis and report (Charmaz, 2014, p. 13). This study employs a constructivist approach to recognize the role of the researcher’s view and the teachers’ accounts in constructing and producing knowledge.

A discussion of researcher reflexivity provides a transparent and explicit explanation of how subjective factors, including the researcher’s personal and political intentions, shape the construction of the research’s design, the development of the research questions, and the choice of data sources and classification schemes (Hughes, Pennington, & Makris, 2012). Chapter Five concludes the report with a reflection on the researcher’s place in the study and the experiences that catalyzed its inception. Charmaz (2005) recommends that grounded theorists should claim “audible voices” in their

writings, which “sparks the reader’s imagined involvement” in the analysis while “illuminating intersubjective worlds” (p. 529). The emphasis on social processes within grounded theory enables researchers “to analyze relationships between human agency and social structure” and to provide an analytic interpretation of participants’ experiences of their socially constructed realities (Charmaz, 2005, p. 508). Chapter Two includes a detailed discussion of the foundational theories used to construct the conceptual framework that explains how teachers’ understanding of value-added evaluation policy shapes how they respond to it. This study presents a constructivist approach to grounded theory that is consistent with the researcher’s stance, accounts for researcher reflexivity, gives an audible voice to teachers, and constructs a grounded, ecologically situated theoretical model that can empower teachers and other practitioners who respond to policy and shape implementation through their daily choices in praxis.

Phenomenological Approach

While several research approaches exist within phenomenology, they all “seek to explore, describe, and analyze the meaning of individual lived experience” (Creswell, 2012, p. 19) and share a focus on exploring how people “make sense of experience and transform experience into consciousness” (Patton, 2015, p. 115). Van Manen (1990) described the purpose of hermeneutical phenomenology: “Phenomenology aims at gaining a deeper understanding of the nature or meaning of our everyday experiences” (p. 9). These approaches review how individuals experience phenomena, interpret those experiences, construct understandings of phenomena with their interpretations, and make sense of their world (Patton, 2015). There is no separate or objective reality from those

interpretations, only the reality constructed from the subjective experience. The Thomas Theorem (1928) stated, "If men define situations as real, they are real in their consequences" (p. 28). This study looks beyond the objective properties of teachers' social settings and considers teachers' meaning making to understand their responses and adaptations in practice to evaluation policy.

While this study borrows from phenomenological approaches, it eschews phenomenological reduction, an analysis method that systematically distills the experiences of the participants to an essence or defining characteristic of the phenomena (Creswell, 2012). In contrast, the study aims to construct a model to incorporate variations in the subjective experiences of the participants. However, two aspects of phenomenology are particularly useful for this study. First, phenomenological interviewing methods complement the sensemaking theories used in the study's conceptual framework, and phenomenology uses in-depth interviewing to describe the meaning of a phenomenon shared by multiple individuals. Marshall and Rossman (2011) explained that phenomenological interviewing "focuses on the deep, lived meanings that events have for individuals" (pp. 148-49) and assumes that their sensemaking of phenomena directs their actions. To construct a theoretical model of how teachers adapt to value-added evaluation, uncovering teachers' beliefs, opinions, and sensemaking of value-added evaluation is necessary.

One element of heuristic inquiry approaches within phenomenology particularly supports this research design because the researcher is a member of the critical reference group of the study. Giorgi (2006) contended that "the researcher has to bracket personal

past knowledge and all other theoretical knowledge ... so that full attention can be given to the instance of the phenomenon” (p. 355). As an elementary education teacher within the Community School District, the researcher brings her experiences of the phenomenon under study to the analysis. While Moustakas (1994) acknowledged that bracketing is rarely accomplished, he described how investigators attempt to set aside their past experiences to allow for a clean perspective of the phenomenon. This study employed Creswell’s (2012) adapted bracketing procedures to include a discussion of the researcher’s standpoint later in this chapter and a description of the researcher’s biases and assumptions in Chapter 1.

Narrative Inquiry Approach

Teachers’ individual characteristics and social contexts cause them to experience policy differently; therefore, semi-structured narrative interviews are used to elicit stories of personal life experiences which in turn reveal each teacher’s unique sensemaking of policy. Collectively, teachers’ narratives also expose “cultural and social patterns through the lens of individual experiences” (Patton, 2005, p. 115). This study borrows from narrative inquiry approaches to generate and analyze stories of personal life experiences through narratives shared in semi-structured interviews. Based in social constructivist theory, this narrative method acknowledges and values teachers’ situated knowledge by remaining committed to local contexts rather than attempting to uncover a single reality.

Genat (2009) claimed “Where representations of the world are highly contested, research becomes a political act,” (p. 105). Narrative inquiry approaches have the potential to advance feminist and critical theory within education research by eliciting the

voice of teachers and validating their construction of meaning through storytelling (Marshall & Rossman, 2011). Furthermore, Denzin and Lincoln (2005) added, “narrative inquiry can advance a social change agenda ... collective stories can form the basis of a social movement. Telling the stories of marginalized people can help to create a public space requiring others to hear what they do not want to hear” (p. 642). According to Mutua and Swadener (2004), critical personal narratives can act as counter narratives that “disrupt and disturb discourse by exposing the complexities and contradictions that exist under official history” and “criticize ‘prevailing’ structures and relationships of power and inequity in a relational context” (p.16). The application of critical narrative analysis methods in this study supports the integrity of the research design and aims to validate the voices of teachers and the meaning they ascribe to their experiences.

Critical Ethnographic Approach

Critical ethnography was developed from a commitment to radical education, but the genre can go beyond the classroom to question historical forces shaping societal patterns (Marshall & Rossman, 2011, p. 26). This study applies a critical ethnographic approach to data collection with the premise that researchers have what Madison (2005) calls “an ethical responsibility to address processes of unfairness or injustice within a particular lived domain” (p. 5). Creswell (2012) explained, “the major components of a critical ethnography include a value-laden orientation, empowering people by giving them more authority, challenging the status quo, and addressing concerns about power and control” (p. 93). The goal of the critical approach is authenticity and credibility of the overall cultural interpretation. This study constructs an emic perspective of teachers using

their own words and filters the findings through the lens of the researcher as a member of the critical reference group. The report provides readers with an understanding of how teachers, as a culture-sharing group, experience and respond to a particular evaluation policy.

Critical Social Theory Approach

Genat (2009) suggests that the search for grand theories is “less urgent in comparison to generating greater understanding of how people can transform their particular life situation for the better” (p. 108). Similarly, Kemmis (2006) advised that research in education without a “critical edge” is inadequate at best, and can become a vehicle for “domesticating students and teachers to conventional forms of schooling” (p. 459). Furthermore, studies designed to improve the efficiency of practices rather than to evaluate their social, cultural, economic, and historical consequences run the risk of reproducing “irrational, unjust and alienating consequences of many existing forms of schooling” (Kemmis, 2006, p. 460). In addition, Kemmis (2006) argues that research work that declines to enter into critical conversations with others and avoids uncomfortable truths cannot bring practitioners closer to confronting the most significant and challenging issues in education and society. Echoing this critical perspective, Kincheloe (1991) recommends that the “‘critical teacher’ exposes the assumptions of existing research orientations, critiques the knowledge base, and through these critiques reveals ideological effects on teachers, schools, and the culture’s view of education” (p. 31).

Denzin and Lincoln (2008) assert that “nothing stands outside representation” (p. 641). Rejecting the notion of neutrality as a sterile environment, free of contamination by researcher and participant bias, critical social researchers can offer alternative spaces where all biases are revealed, acknowledged, and claimed in a negotiated construction of knowledge. Postmodern approaches embrace the variations among contexts and the complexities and complications of lived lives (Clarke, 2003). Feminist methodologies assert that reflexivity in qualitative research is a strategy for replacing pretenses of objectivity with subjectivity (Marshall & Rossman, 2011). Constructivist grounded theories in particular are formed to validate biases as the source of diversity among the perspectives, relationships, and constructed meanings gathered together in the research process. By giving voice to the unheard, feminist perspectives validate the subjective experiences of teachers and “uncover cultural and institutional sources and forces of oppression” (Marshall, 1997, p. 12).

Researcher’s Standpoint

In addition to meeting the standard Marshall and Rossman suggested for compelling research, the selected methods reflect an ontological and axiological integrity with the philosophical standpoint of the author of this study. As a critical social scientist, the researcher’s frameworks borrow heavily from postmodern critical social and cultural theories and social constructivism. These ontological perspectives recognize multiple realities based on power and identity struggles between individuals and groups. These realities can become known through a study of social structures and discourses and where control over knowledge resides.

The researcher also embraces the transformative ambitions of Freire and Gramsci to develop an informed, critical perspective of education policy. Exploitative relationships with knowledge can create hegemonic narratives used to control society, but they can also serve as catalysts for action. Freire contended that a radical education begins with people's current reading of the world and enables them to "change their relationships to knowledge" (Allman & Wallis, 1992, p. 13). As a social constructivist, the researcher believes individuals create shared meanings through their interactions and those meanings become their reality. Identity and agency are social constructions; an individual's process for making sense of his or her experiences and acting on them cannot be done in isolation. Sociologist Bourdieu (2000) argued that agency "is of a socialized body, investing in its practice socially constructed organizing principles that are acquired in the course of a situated and dated social experience" (pp. 136-137). From a critical theory perspective, the researcher has a responsibility to deconstruct the role of power in defining realities, to construct knowledge as a social process, and to democratize access to knowledge in the pursuit of changing the dynamics of power and knowledge that marginalize teachers. Using a postmodern approach to grounded theory, this qualitative study is designed to establish ontological, epistemological, and axiological integrity between the philosophical beliefs of the researcher, paradigms of knowledge, and research methods.

In summary, several complementary methods of empirical research design support the objectives of this research. This constructivist, grounded theory study applies critical narrative and ethnographic inquiry and analysis methods to produce

epistemologically valid, empirical evidence through a range of sources including interviews, memos, and recordings (Hughes et al., 2012). By naming the inherent subjectivity of traditional research and the relationship between power and knowledge, educators and researchers who reside within participant member groups can critically review and improve research design elements that have excluded teachers' lived experiences in the construction of knowledge. The "audible voice" of the researcher and participants that are central to the design of this study support an epistemological claim of providing a truth of this particular participant group's experiences at a particular moment (Genat, 2009, p. 114). This study has produced an evocative report for evaluation policy stakeholders that "enables an empathetic understanding of the experience" of teachers and invites readers to "recognize similar experiences of their own within the text" (Genat, 2009, p. 114). The next section describes the methods that were used to collect data that addressed the central question of this study.

Description of Methods

The purpose of this study is to explain teachers' responses to value-added evaluation policies through their reflections and narratives as empirical research for understanding policy dynamics and teachers' role in the implementation process. This study contributes to the limited qualitative literature on teachers' responses to value-added models and professional evaluation policies by answering the central question: What theoretical model can explain the process of how teachers have adapted their practices in response to value-added evaluation?

Methods

This section begins by describing the data collection approach, Community School District data, the participant selection process and ethical considerations for participants. The section continues with a description of data collection and data analysis methods. A discussion of the criteria used to evaluate quality of the research concludes the section.

Data collection approach. The study draws from phenomenological, narrative and feminist approaches to collect data. Phenomenological research seeks to uncover the lived experiences of a group of people who share a common experience, or phenomenon (Creswell, 2012). The purpose of phenomenological interviewing is to describe the meaning of an experience, or phenomenon, that a group of people share (Marshall & Rossman, 2011). Data is collected through in-depth interviews to capture and illustrate how teachers experience value-added evaluation (Patton, 2015). Polkinghorne (2007, 2010) suggested the value of using interviews and personal stories in education research. This study adopts elements of the narrative inquiry approach to learn how teachers have adapted to a particular value-added evaluation policy, to illuminate how teachers make sense of that policy and to explain the process of how teachers have adapted to policy implementation. Semi-structured narrative interviews are used to collect stories from teachers and represent their experiences in their own words. Narrative interviews are ideal for this study because personal stories can help researchers understand adaptations of groups through the interpretation of “individuals’ lived experiences” (Marshall & Rossman, 2011, p. 153). This approach is an effective method for data collection because

it provided teachers a platform to explain how they adapt to evaluation reforms. DiPlacito (2011) suggested that “with a focus on the articulation of personal stories, the interview approach serves as an admirable method of data collection” (p. 13) in the study of women and their life experiences. This inquiry incorporates feminist approaches to resist the notion that personal narratives are primarily useful for understanding the impact of social structures on individual’s lives. Instead, participants here are seen as “social actors in their own right” (p. 655), and the research looks to understand what meanings the teachers assigned to the conditions of their professional and personal lives (Chase, 2005). With a narrative structure and feminist aims, the interviews are employed to actively involve the participants in the telling of their personal experiences and localized perspectives (Creswell, 2012; Edman, 2005).

District data. In the spring of 2018, Community School District (CSD) reported serving more than 10,000 students and employing more than 800 state-certified teachers. For the 2016-17 school year, the district reported an annual budget in excess of \$90 million with per pupil expenditures of over \$8,000. For the same school year, the Tennessee Department of Education (2017) reported a graduation rate of over 90% for the district.

The students of CSD predominantly self-reported as “White,” less than 5% as “Hispanic,” and less than 5% of students split between “African American”, “Asian”, “Native Hawaiian/Pacific Islander” and multiracial categories. Also in 2016-17, the district served over 200 English language learner students. Based on October enrollment

in 2016, nearly 30% of students in the district were identified as “economically disadvantaged,” (Tennessee Department of Education, 2017).

Participant selection. Theoretical sampling (Creswell, 2012; Glaser & Strauss, 1968; Lempert, 2007; Maxwell, 2012; Miles & Huberman, 1994) is used to select participants for the study. Theoretical sampling requires data collection until the data support theoretical saturation. To be considered for the study, participants must have previously been evaluated or are currently being evaluated as an elementary school teacher using TVAAS in the Community School District, and they must be employed by the district at the time of the interview. Participants are identified using existing professional and personal networks within the Community School District where the researcher is employed. The Middle Tennessee State University’s Institutional Review Board (IRB) and the Assistant Director of Schools/Director of Human Resources provided permission to conduct the study. Participants are contacted and invited to join the study through email. Interested candidates receive additional information including informed consent documents via an email with a link to a Qualtrics survey. Phone interviews are arranged once participants provide informed consent. Through constant comparative analysis, additional participants are recruited using these protocols to reach theoretical saturation.

Ethical considerations. All participants are interviewed by phone. Participants are informed that they can leave the study at any time and can request that all data related to the participant be withdrawn from the study documents and archives. Interviews are audio-recorded and the interviewer transcribes all of the recordings. Participants are

informed that member checking techniques are used. Responses may create vulnerabilities for participants and may affect responses; therefore, the participants are guaranteed access to the information printed about them, and they are informed that they have the option to delete any part of the transcript after reviewing it. In addition, findings are shared with participants before public dissemination. Participants are advised that participation is confidential. Due to the narrative nature of the interviews, not even pseudonyms are used to protect the anonymity of participants. Any mentions of persons and cities are kept confidential between researcher and participant. Data is stored and backed up on the researcher's personal devices. All hard copies are stored in the researcher's home. In addition, none of the documents are stored on school property or devices. Participants have exclusive access to their transcripts during the study.

Data collection methods. The interview protocol is designed to discover the experiences and perceptions of value-added evaluation policy among teachers from the perspective of the individuals who have been evaluated with this policy. This grounded theory study aims to give voice to those whom the policy was intended to influence; therefore, a semi-structured narrative interview approach has been selected to elicit information about the perceptions and agency of the participants. A value-added evaluation system is used in the Community School District as an intervention to impact the behavior of classroom teachers, as objects of policy. This study uses narrative interviewing to serve the participants by giving voice to their stories of adaptation. One-on-one, semi-structured interviews ask participants about their understandings of value-added evaluation and their adaptations to being evaluated using value-added methods.

The initial question asks participants to share their stories of how they became a teacher. Life histories in this study are not used to answer the research question or to generalize experiences. Instead, they are integrated into the protocol to express empathy and caring for the participants (Marshall & Rossman, 2011). To generate rich descriptions of participants' perceptions and experiences, their own words are used to generate probing follow-up questions (Roulston, 2010). To minimize the effect of changes in the context of the study on participants' responses, the researcher conducted all interviews within a three-month period.

Data analysis methods. This study applies several grounded theory tools and strategies to support an “emergent intuitive” design for data analysis (Marshall & Rossman, 2011, p. 209). Analysis relies on constant comparative analysis, memoing, and coding to simultaneously collect and analyze the data. As explained above, theoretical sampling is used to saturate the theoretical categories that emerge from the data and to construct a substantive theory (Charmaz, 2014) Analysis begins with a constant comparative analysis to find commonalities and differences within the transcripts, interview notes, and memos. The empirical evidence from narratives of the participants is compared to theoretical coding categories that emerge throughout the simultaneous collection and analysis of data. (Creswell, 2012; Glaser, 1965; Schwandt, 2007). Each incident is compared with other incidents. Through reiterative cycles of comparison, conceptually similar incidents are grouped together under broader descriptive categories which are refined to “crystallize participants’ experience” (Charmaz, 2014, p. 133; J. M. Corbin & Strauss, 2008, p. 73). Both memoing and coding practices are essential tools for

developing the theoretical coding categories. Miles and Huberman described analytical memos in grounded theory as “one of the most useful and powerful sense-making tools at hand” (1994, p. 72). Memos are the “space and place” (p. 163) where ongoing comparisons are made within the empirical evidence that is collected in the interviews (Charmaz, 2014). Categories, concepts, and emerging theories are recorded through memo-writing, interpretations, and analyses of initial codes (Charmaz, 2014; Corbin & Strauss, 2008, p. 118; Glaser, 1978; Lempert, 2007; Marshall & Rossman, 2011, p. 213; Miles & Huberman, 1994; Saldaña, 2012, p. 41). Lempert (2007) described memo writing as “the distillation process, through which the researcher transforms data into theory” (p. 245). Through the use of memos, the participants, their narratives, and the emerging processes are systematically reflected on to document a “path of theory construction” (Charmaz, 2014, p. 164).

Charmaz (2014) suggested that coding in grounded theory consists of at least two stages, initial (or open) and focused coding. Initial coding for this study begins soon after the first interview and continues with constant comparative methods to ask what the comparisons within the data suggest about the narratives of teachers (Charmaz, 2014; Corbin & Strauss, 2008; Schwandt, 2007). Heuristic initial codes remain close to the actions and processes in the data, not the topics, to avoid making assumptions or imposing a priori theories too early (Charmaz, 2014; Saldaña, 2012). The codes serve as the starting point to develop ideas for further exploration (Saldaña, 2012). A list of all emerging codes from this initial stage is found in Appendix A. Next, in the focused coding stage, the researcher revisits the initial codes to decide which best support an

analysis of the entire data set (Charmaz, 2014). Through focused coding, (an update of axial coding in classic grounded theory) codes that best describe the data and reveal gaps in them are identified (Charmaz, 2014; Saldaña, 2012). The most salient codes are clustered together according to emic, situated categories and then assigned tentative codes to represent their relationships and connections (Carley, 1993; Maxwell, 2012; Saldaña, 2012). In the final coding stage, the focused codes are analyzed according to how they relate to one another to develop a theoretical code that “functions like an umbrella that covers and accounts for all other codes and categories” (Saldaña, 2012, p. 223). Additional participants are interviewed to explore and test the adequacy of the theoretical category until “theoretical sufficiency” (Charmaz, 2014, p. 215) or “saturation” is reached (Saldaña, 2012, p. 227; Schwandt, 2007, p. 110). Drawing from Corbin and Strauss’ (2008) framework for analyzing data for process, the categories are analyzed to develop a theoretical model that explains the participants’ adaptation to value-added evaluation. This process model represents the sequence of actions, interactions, and emotions that change in response to combinations of contexts, conditions, events, and interactions based on the researcher’s interpretation of the data (Corbin & Strauss, 2008).

Guba’s Criteria for Quality of Research

The next section describes the criteria and measures that are selected to establish the quality of this inquiry through the lenses of trustworthiness and adequacy of the research. Guba (1981) called for a balance between the rigor of quantitative methods and the relevance of qualitative approaches by offering four criteria to judge the

“trustworthiness” of research, including credibility, transferability, dependability and confirmability (pp. 79-80).

Credibility. The study’s methodology incorporates several measures to address credibility issues. As a classroom teacher in her seventh year at the initial research site, the researcher is familiar with the setting and many of the participants. Interviews and observations occur over several months and provided sufficient time for data collection at the site. Prolonged engagement and persistent observation at the research site support the credibility of the data (Guba, 1981). However, the researcher’s familiarity and membership with the participants’ culture-sharing group, can create a threat of local bias which would have influenced the interpretations (1981, p. 84). Peer debriefing and researcher journals mitigate the limitations posed by familiarity. Peer debriefing with fellow doctoral students and dissertation committee members provide the researcher opportunities to detach from the research site and test grounded theories as they emerge and evolve. In addition, the researcher’s peers challenge her thinking with questions that lead to new insights. Journals and field activities provide a place to reflect and challenge the researcher’s role in collecting data and the meaning she ascribes to them. Journals are also a place to wrestle with feedback and questioning from debriefings (Guba, 1981). Through research journals, the researcher examines how her presence affects phenomena and her observations and interpretations of it (Patton, 2002). Another way to determine the credibility of a naturalistic inquiry requires an assessment of how the data represent the meaning intended by the participants and researcher. This was beyond the scope of this study.

Credibility can be improved with several triangulation measures (Guba, 1981; Shenton, 2004). By including a wide range of informants, the researcher is able to verify individual perspectives against each other to paint a rich landscape of identity, attitudes, and agency (2004). In addition, interviewing teachers from seven different elementary schools with collective experiences from 11 different schools within the district reduces the effect of specific factors peculiar to a single school. Similar findings emerge from multiple schools which enhances the credibility of the study (Shenton, 2004). With this grounded theory study, a negative case analysis is used to refine emerging theories until one unifying theoretical model is able to address all of the cases within the data set. By collecting referential adequacy materials including audio recordings of interviews, interview notes, journals, and memos, the researcher tests findings and interpretations against the archive of data (Guba, 1981). One of the most critical measures for improving credibility is conducting member checks with the participants (Guba, 1981; Shenton, 2004). The validity of the data is tested by asking follow up questions during the course of the interviews and by soliciting feedback on the accuracy of transcripts of interview dialogues from the participants (Schwandt, 2007).

Transferability. This naturalistic inquiry does not attempt to make generalizations or transferability inferences because the phenomena described are bound to the historical moment and context of the schools where the data are collected (Guba, 1981; Shenton, 2004). Instead, the aim of the study is to collect rich, descriptive data. This will provide enough description of context and boundary of the study so that the reader can test the “degree of fittingness” (Guba, 1981, p. 81). The final report of this

study aims to provide a basis for subsequent research on organizational theory, agency, and teacher evaluation policies. This is one isolated project; understanding teacher perceptions and agency in response to value-added methods of evaluation can evolve only through a collection of studies (Shenton, 2004, p. 71).

Dependability. The researcher in this qualitative study is also the instrument of data collection. Guba (1981) argued that humans as research instruments bring instability to findings due to our changing nature. Yet instruments must produce stable results if results are to be dependable. Reliability is a precondition for validity (Guba, 1981). While invariance is impossible, tracking variance is possible. Some researchers suggested that processes must be made more public to improve the rigor and defensibility of qualitative research (Anfara, Brown, & Mangione, 2002). A transparency of the processes within this study provides readers with enough details to repeat the work described here.

Confirmability. With social science inquiry, the researcher's value systems and predispositions inevitably bring bias to design (Shenton, 2004). Neutrality is neither possible nor is it the goal of a naturalistic study. Instead, Guba (1981) insisted that naturalistic inquirers work toward confirmability by "practicing reflexivity" (p. 87). Therefore, in addition to the measures described above, this report also explicitly addresses the researcher's beliefs and assumptions in the "Rationale for Methods" section.

Charmaz's Criteria for Quality of Grounded Theory Research

Charmaz (2005) argued that "a grounded theory born from reasoned reflections and principled convictions that conveys a reality makes a substantive contribution" (p.

529). This study aims to meet Charmaz's (2005) criteria for judging the quality of grounded theory research include credibility and originality. As described above, the methods of this study incorporate multiple measures for establishing credibility with the reader including prolonged engagement, peer debriefing, triangulation, and member checks. The study also provides a new conceptual framework with an original application of institutional and sensemaking theories to provide an explanation of the process teachers' experience when responding to evaluation policy.

This chapter presented the impetus for the design of this study and a rationale for a constructivist grounded theory approach. A detailed discussion of the design was followed by a description of methods. The chapter concluded with an assessment of criteria for the quality of the research. Chapter 4 presents the study's findings, which include the core themes generated from the open and selective codes and the theoretical model generated to respond to the central research question. Chapter 5 provides a response to the empirical research questions and situates the theoretical model. Discussions of the implications for practice and policy as well as recommendations for further research are also included followed by the researcher's reflection and closing remarks.

CHAPTER IV

RESULTS AND ANALYSIS

Context of the Study

In January 2010, Tennessee General Assembly passed a bill requiring school districts to use Tennessee's Value-Added Assessment System (TVAAS) and other student outcome measures as a significant component of teacher evaluations starting in the 2011-12 school year (Piro et al., 2011). The law mandated school districts in Tennessee to use student scores on state standardized tests to calculate teacher effectiveness scores. Observation scores from the TEAMTN rubric and other qualitative measures accounted for 50% of evaluation criteria using a five-point scale (1-5). The remaining half of evaluation scores was determined by two categories of student performance on state standardized tests. Fifteen percent was calculated from measures of student achievement based on student proficiency. The remaining 35% was calculated from measures of student growth. According to the Tennessee Value-Added Assessment System (TVAAS), these measures were used to represent the impact of a teacher on his or her students' growth in achievement scores over the prior school year. This impact, called value-added, is quantified using the value-added methods of TVAAS. At the time of this study, a teacher's TVAAS data were used to determine his or her value-added score that contributed 35% of his or her overall effectiveness score. This study considered the use of value-added scores in teacher evaluation in Tennessee as a critical case to illuminate the dynamics of policy implementation by exploring how teachers understood

value-added evaluation policy and how their understandings of policy shaped their choice of practice in classrooms and schools.

Purpose of Study

The purpose of this study was to explain teachers' responses to value-added evaluation policies through their reflections and narratives as empirical data for understanding policy dynamics and teachers' role in the implementation process. The data collection for the study was guided by the following research question:

- What theoretical model can explain how teachers have adapted their practices in response to value-added evaluation?

From this central question, two empirical questions followed:

- What are the lived experiences of teachers who have been evaluated by TVAAS?
- What practices do teachers choose in response to TVAAS evaluation policy?

This chapter discusses the findings and analysis of data provided by the 19 participants in the study. The central research question guided the study to develop a theoretical model to explain how teachers adapt to evaluation policy through choices in practice. The chapter begins with summaries of how the data were collected, the demographics of the participants, and the methodology used for analyzing the data. Following these summaries, the study's findings are organized into three sections: analysis of the

interview questions, analysis of the four core themes and synthesis of the theoretical model.

Data Collection

The results presented in this chapter were drawn from single, semi-structured interviews with 19 elementary education teachers from the Community School District, a K-12 school district in Tennessee, a state that implemented a value-added component to teacher evaluations in 2011. After participants provided informed consent, a Qualtrics survey was administered to collect demographic information and interviews were scheduled. Elementary school teachers who were being evaluated using TVAAS at the time of the interview or had been evaluated using TVAAS prior to the interview were eligible to participate in the study. All participants were interviewed by phone for consistency in data collection and to reduce potential bias in participation based on logistical challenges of meeting face-to-face. Through constant comparative analysis, additional participants were recruited to reach theoretical saturation. To minimize the effect of changes in the context of the study on participants' responses, the researcher conducted all interviews within a three-month period between April and June of 2017. This time period began after state standardized testing and ended before scores were reported. All phone interviews were audio recorded and transcribed for analysis.

Demographics of Participants

To be considered for the study, participants were required to meet several criteria. Participants must have been evaluated previously or were being evaluated currently as an elementary education teacher using the TVAAS scores from their own students (versus school-wide TVAAS scores) in the Community School District at the time of the interview. From the selected research site, theoretical sampling (Corbin & Strauss, 2008; Maxwell, 2012; Miles & Huberman, 1994) was used to select participants for the study. Potential participants were identified using existing professional and personal networks within Community School District where the researcher is employed as an elementary education teacher.

Survey Questions

Qualtrics software was used to collect information about participants' professional histories, including years of elementary education teaching experience, current and past employment within the Community School District, gender, evaluation status, and years of TVAAS evaluation.

Participants' Level of Elementary Education Teaching Experience

Table 4.1 shows the participants' number of years of teaching experience in elementary schools over the course of their careers and the number of years of teaching experience at their current school. Nearly 80% of participants had over five years of elementary education teaching experience. Participants with over 10 years of experience comprised 52% of the sample. These teachers had a minimum of four years of teaching

experience before the TEAMTN evaluation system (including the TVAAS component) was adopted by the Community School District.

Table 4.1. Participants' level of elementary education teaching experience

Years	Career experience in elementary schools		Experience at current school/position	
	Number of participants	Percentage of participants	Number of participants	Percentage of participants
0-5 years	4	21%	9	47%
6-10 years	5	26%	5	26%
11-15 years	5	26%	5	26%
16-20 years	2	11%	0	0%
Over 20 years	3	16%	0	0%
Total	19		19	

Participants' Past and Current Employment by School within the Community

School District

Table 4.2 illustrates the distribution of participants' past and current employment within the Community School District over the course of their teaching careers.

Collectively, the participants have taught at 11 of the district's 14 elementary schools. At the time of the interviews, 17 participants taught at seven different elementary schools, and two of the participants had moved out of the classroom into district-wide coaching positions. Letters were used in place of school names to maintain confidentiality.

Table 4.2. Participants' past and current employment by school within the Community School District

Community School District elementary school	Number of participants formerly employed	Percentage of participants formerly employed	Number of participants currently employed	Percentage of participants currently employed
A	1	5%		
B	4	21%		
C	2	11%		
D	2	11%		
E	0	0%		
F	2	11%	1	5%
G	0	0%	1	5%
H	3	16%		
I	4	21%		
J	3	16%		
K	8	42%	3	16%
L	0	0%	1	5%
M	2	11%	2	11%
N	5	26%	5	26%
O			4	21%
Central Office	2	11%	2	11%
Total			19	100%

Participants' Gender.

Table 4.3 displays the gender from the participant pool: 84% female and 16% male. The gender distribution of elementary school teachers in the Community School District is 91% female and 9% male.

Table 4.3. Participants' gender

Gender	Number of participants	Percentage of participants
Female	16	84%
Male	3	16%
Total	19	

Participants' Evaluation Status

Through the Qualtrics survey and follow-up interview questions, participants reported if their current evaluations would include a TVAAS component calculated from their roster of students, not school-wide TVAAS scores. Table 4.4 presents the evaluation status of the participant group at the time of the interviews by the number and percentage of participants. The sample pool split with 12 participants (63%) reporting “Yes” and 7 participants (37%) reporting “No.” It is worth noting that testing in grades 3-8 was cancelled for the 2015-16 school year in Tennessee when Measurement, Inc. failed to provide necessary printed testing materials to schools (Rainwater, 2016). Three participants initially indicated that they were “Unsure” if their evaluations at the time of the interview would include TVAAS scores. When asked about their uncertainty during interviews, two of the respondents indicated that issues with TNReady standardized testing in the spring of 2016 left them uncertain if their current evaluations would include TVAAS scores. Their responses were revised to “Yes” because they confirmed that their evaluations included TVAAS scores based on their roster of students under normal circumstances. The interview with the third participant to select “Unsure” revealed that the participant had recently transferred to a district-wide position that was not evaluated by their own students' TVAAS scores. This participant's response was revised to “No.”

Table 4.4. Participants' evaluation status

Currently evaluated by students' growth scores	Number of participants	Percentage of participants
Yes	12	63%
No	7	37%
Total	19	

Years of Participants' TVAAS Evaluation

Participants were asked in which academic years did their annual evaluation include a TVAAS score based upon the growth scores of their students as a component of the TEAM evaluation system. Table 4.5 presents the number of teachers and the percentage of teachers who were evaluated with their students' growth scores for each academic year since the implementation of TVAAS evaluation policy in the Community School District until the interview period. For the participant group, the peak years of TVAAS evaluation were 2011-12 at 63% of participants and 2012-13 and 2014-15 at 68% each. Since 2014-15, the percentages have fallen to 47% in 2015-16 and 53% in 2016-17.

Table 4.5. Years of participants' TVAAS evaluation

Academic years evaluated with TVAAS	Number of interviewed teachers	Percentage of interviewed teachers
2010-11	12	63%
2011-12	13	68%
2012-13	13	68%
2013-14	12	63%
2014-15	13	68%
2015-16	9	47%
2016-17	10	53%

Methodology of Data Analysis

Goal of Analysis

The interview process produced several layers of data, including interview transcriptions, detailed notes during interviews, and researcher memos. The goal of data analysis was to derive core themes to construct a theoretical model to explain the process teachers experienced in choosing practices in response to value-added evaluation policy. While interview notes and memos informed the analysis, the core themes were generated from the 19 participants' words and experiences recorded in the interview transcripts.

Coding Procedures

Following the grounded theory coding procedures, the researcher analyzed three data sources, including interview transcriptions, interview notes, and researcher memos in three stages: open, selective, and theoretical (Glaser, 1978; Urquhart, 2013). Based on interview notes and memos, the researcher assigned open codes to responses soon after the first interview. The researcher also assigned codes to participants' responses to link the responses with the interview questions that evoked them and enable frequency counts of specific codes by question later in the analysis. In addition, the researcher assigned multiple, heuristic codes to participants' responses that remained close to the practices and experiences in the transcripts, not the research topics, to avoid making assumptions or imposing a priori theories too early (Charmaz, 2014; Saldaña, 2012). The codes served as the starting point to develop ideas for further exploration (Saldaña, 2012). As this initial coding process was re-iterated with additional interviews and subsequent notes and memos, the researcher added, consolidated, and eventually grouped other open codes into

several categories in the first selective coding stage using the constant comparative method (Glaser & Strauss, 1968). In this selective, or focused, coding stage, the researcher immersed herself in the data through repeated readings and by revisiting the initial codes to decide which could best support a cohesive analysis of the data in their entirety (Charmaz, 2014). Through focused coding, (an update of axial coding in classic grounded theory) the researcher identified codes that could account for both convergences and divergences in the data (Charmaz, 2014; Saldaña, 2012). These most salient codes were clustered together according to emic, situated categories, and four core themes representing their relationships and connections were developed (Carley, 1993; Maxwell, 1996; Saldaña, 2012) Next, the researcher analyzed the relationships and dynamics among the four core themes to develop a working theory that “functions like an umbrella that covers and accounts for all other codes and categories” (Saldaña, 2012, p. 223). To test the adequacy of the model, the researcher recruited additional participants based upon their years of teaching experience, the schools where they had taught, their years of TVAAS evaluation, and their gender until “theoretical sufficiency” (Charmaz, 2014, p. 215) or “saturation” was reached with 19 interviews (Saldaña, 2012, p. 227; Schwandt, 2007, p. 110). The themes that integrated the findings across the 19 interviews were maintained while those that appeared inconsistent were re-examined to modify and adapt the working theory as it developed over the course of the interviews. By consolidating and eliminating recurring themes, the researcher identified four core themes from the experiences and beliefs of the participants to ground a working theory.

Adaptations in coding. One example of this adaptive coding process came early in the constant comparative analysis of interviews and notes. The study was designed to explain the relationships between teachers' understanding of policy and their choices in practice. After the first several interviews, it became evident that participants' understandings of policy were diverse while many commonalities of institutional and teacher practices surfaced across different schools within the district. The relationship between participants' explicit understanding of policy and practices appeared less significant than initially anticipated. Instead, the experiences of participants with institutional practices related to value-added evaluation surfaced as subjective phenomena that formed individual teacher choices in their practices. These phenomena significantly shaped the codes that were used to describe the data.

Findings and Analysis

The summary of findings is organized into three sections: analysis of the interview questions, analysis of the four core themes, and synthesis of the theoretical model. In the first two sections, the words of the participants illustrate the categories and core themes constructed from the interviews, along with the researcher's interview notes and memos. The third section integrates the four core themes constructed from the codes and categories to generate a theoretical model that explains the process teachers experience in choosing practices in response to value-added evaluation policy.

Analysis of the Interview Questions

The purpose of this constructivist grounded theory study (Charmaz, 2014; Corbin & Strauss, 2008) was to build a theoretical model to explain how teachers have adapted their practices in response to value-added evaluation. The open-ended interview questions were designed to create a space for teachers to communicate their meanings of the policy. The questions also provided an opportunity for teachers to reflect on their choices in practice within the social context of their school and teams. The interview protocol included questions related to participants' professional histories and current beliefs about teaching and evaluation. These were included in the interviews to establish a rapport between the interviewer and participants. The list below includes interview questions relevant to the research question, "what theoretical model can explain how teachers have adapted their practices in response to value-added evaluation?"

- In your opinion, what is the best use of teacher evaluations?
- In your opinion, what is the intended purpose of the value-added component in teacher evaluations? (as part of the Tennessee Educator Acceleration Model)
- Does the way you're evaluated shape your teaching? Why or why not? If yes; please tell me about how your anticipated value-added score(s) shape(s) your teaching practices.
- Please describe interactions you have had with colleagues about value-added scores and the ways, if any, that your team adapts or organizes instruction in anticipation of value-added scores.

- Please describe any interactions you have had with principals about value-added scores and the ways you are aware of, if any, that your principal has implemented team or school-wide practices related to your school's value-added scores.

Participants' responses to each question were assigned heuristic codes in the initial phase and subsequently consolidated and adapted into recurring categories. Tables were produced with the recurring categories on one axis and individual teachers on the other to generate frequency counts of themes by participant. This section presents a summary of the analysis for the interview questions. Due to the narrative interview design, the questions elicited detailed accounts of episodes where their peers and supervisors were present. Collections of narratives by a single participant created vulnerabilities for confidentiality that could have revealed that participant's identity. Therefore, to protect the confidentiality of all participants, all responses were reported anonymously in the study.

Interview questions regarding best use of evaluations. This question asked participants to reflect on their beliefs and values regarding evaluations in general. Teacher Improvement was most widely cited as the best use of teacher evaluations. A related category, Constructive Criticism, was also noted by more than a quarter of the participants. Table 5.1 presents the categories and frequency counts of participants for each.

Table 5.1. Participants' responses regarding beliefs of the best use of evaluations

Question Specific Categories	Number of Teachers	Percentage of Teachers
Teacher Improvement	11	58%
Constructive Criticism/Feedback	5	26%
Teacher Accountability	4	21%
Teacher Reflection	3	16%
Teacher Recognition	1	5%

Several quotations highlight the five categories, each from a unique participant within each category.

Teacher improvement.

“I think teacher evaluations should be used to help the teacher to improve.”

“I think that's the best use of teacher evaluations is to help a teacher to improve his or her craft.”

“Probably professional development. You know, help guide where [teachers] have weaknesses.”

Constructive criticism/feedback.

“I think the best use of evaluations, that would be to be able to take the constructive criticism from the evaluator, and better yourself.”

“I think for me the best use of my teacher evaluations is to reflect on what I need, what are the constructive criticisms.”

Accountability.

“I also think a level, kind of like a plumb line, of expectations has to be set at some point for certain personalities that won't make changes unless there's more weight to like what the product is at the end of the day from your teaching what has been seen. I guess I know I realize it's an accountability piece. And I guess to some extent I think that probably for some teachers that's necessary.”

Teacher reflection.

“I would like to see it continue where you get you use your evaluation to look at how am I doing as a teacher or the methods I'm using, are they working well? Or perhaps I need to change it up? Those kinds of things. That's very effective.”

Teacher recognition.

“I always liked that as an affirmation that I have done a good job.”

Interview questions regarding policy intent. This question was presented to teachers to reveal their understanding (sensemaking) of value-added evaluation policy. The teachers were asked what they believed to be the purpose of value-added evaluation.

The specific responses from teachers were varied, but several common categories surfaced from the initial codes. Table 5.2 displays these categories along with frequency counts for each. More than half of the teachers identified teacher improvement as the intended purpose of value-added evaluation. Approximately half of the participant sample identified growth of student achievement and accountability of teachers as the intention of the policy. Given the open-ended nature of the question, it is interesting to note that over 40% of participants responded with issues related to the policy without prompting. Several teachers expressed a variety of misuses of the TVAAS scores against teachers including blaming, punishing, and ranking.

Table 5.2. Participants' responses regarding policy intent

Question Specific Categories	Number of Teachers	Percentage of Teachers
Teacher Improvement	11	58%
Growth (student achievement)	9	47%
Teacher Accountability	8	42%
Intent Uncertain or Unknown	4	21%
Determine Teacher Effectiveness	3	16%
Objectivity to Evaluation	3	16%

Below are quotations from teacher responses to illustrate some of the categories the researcher identified. Each quote is from a unique participant.

Teacher improvement.

“I think is it is intended to have them use more differentiated instruction in the classroom, you’re not just teaching one lesson, all the students as one-size-fits-all. You’re looking at where students are when they come in to the lesson, and then you’re differentiating how you teach in looking at whether your teaching is creating growth, and if it's not, you need to change what you are doing constantly.”

“[It’s] also intended for the teacher to be able to use the data to improve based on student performance whether they've missed the curriculum or maybe their classroom management, lots of things to look at.”

“I think that’s originally what it was supposed to be for, for the teacher to look at, to see how to improve and where to improve and then somehow the state has come in and now saying, ‘now we're gonna hold you accountable for the scores.’ They're not there to help you they're there to pretty much shame you.”

“Some people see it as a punishment, some people see it as a motivator, definitely like, just because they want to I guess to be better ... it also intimidates people and stresses them out.”

Growth.

“To see the effect you had in your students, to see have they grown, the amount of levels that they should have grown, or what is expected for them to grow.”

“I think that it’s intended purpose is to check to make sure the students are growing from year-to-year.”

“It's a very basic level to ensure growth of students or to monitor growth of students. And then to attach that to, you know, evaluating whether a teacher is more or less effective.”

Accountability.

“I think the teacher should be held accountable for what their students learn. Because that's kind of our job. Right?”

“I think it's for checks and balances. Because I mean it's a nice thought to think we're all good teachers. It's a nice thought that we're all doing what we're supposed to be doing. That we're all teaching the standards and that the kids under our care are all learning what they're supposed to be learning. That's a really nice thought. But we all know, as teachers, and they all know if they've ever been in a school, that there are teachers that do not do what they are supposed to be doing.”

“You know they want teachers to feel accountable, and well at least that's what I'm assuming.”

Intent is uncertain or unknown.

“I don't know really. I guess to ensure that they're teaching what needs to be taught? I don't know. I never really thought about that.”

Teacher effectiveness.

“I think the purpose is to use the scores as a way to help teachers evaluate are they being effective in the classroom. I don't know that that is what's happening.”

Brings objectivity to evaluation.

“I think for years and years they relied on the principal of each school just to [ask], ‘how are your teachers doing?’ ... and then districts probably realized we need something...that can be more of an equalizer from school to school, teacher to teacher, school to school, and even district to district. And so they looked around, here it was. Here was this formula that a professor at UT came up with, here's something we can use to make that happen.”

Interview questions related to classroom practice. This question and its follow up were intended to prompt teachers to reflect on their classroom practices in response to value-added evaluation. The teachers were asked if and how they adapt their teaching

practices in anticipation of value-added scores at the end of the year. Four categories were constructed from the open codes, including test preparation, planning, instruction, and curriculum, Table 5.3 displays these categories along with frequency counts for each. Notably, among all participant responses, all but one teacher indicated some adjustment in practices in anticipation of value-added scores. Teachers shared a range of practices related to test preparation, including review of test content, test format, teaching testing strategies, conducting practice tests, and building stamina of students to prepare for standardized test conditions. Teachers who indicated adaptations in instruction most often included individualized instruction for students to improve their growth.

Table 5.3. Participants' responses regarding adaptations in practice

Question Specific Categories	Number of Teachers	Percentage of Teachers
Test Preparation	14	74%
Planning	11	58%
Instruction	7	37%
Curriculum	4	21%

Several quotations highlight the four categories, each from a unique participant within each category.

Test preparation.

“You are not only teaching your content but you are also teaching test taking strategies. You are doing practice that is going to be like the test ... You just

make questions that go with what you're doing that are a similar set up. So if they're practicing, they're not having to take time out, so if they're doing a test for practice, no, you just incorporate those kinds of questioning, that kind of questioning, that kind of thing to make sure they're in there."

"We did the practice test the state put out I guess. And then we did it in like in one sitting. Practiced, because the hardest part with my kids was not the actual math, it was it was the 'you can't get up, can't talk, you can't disturb neighbor, you can't tap, you can't sing. No, I'm sorry you can't get up and go blow your nose or sharpen your pencil or go around the room and ask if anybody has an eraser you can use.' That was harder for them than anything."

Planning.

"The first year my kids struggled with life science, so I hit my science really hard the next year."

"You know you are always planning with that final test in mind. Everything that you do if you're testing, you are planning with that final test in mind."

Instruction.

"I think that testing changes at least the way that I have to teach because ... you have all year long to teach your materials, but you really only have until that test day to teach it. You know? Because if you don't teach it by whatever day they're

testing on, they're not going to have that material. So a lot of times I feel that it's like a push to get things done because I have to get it done by a certain time.”

“Several of us wanted to teach small groups all the time, so it looks more like a rotation style going on within the classroom. So we adjusted to that, and adjusted our TAs [teaching assistants] according to how we did them so that we could maximize the teaching with the students, and then we could focus with students on what their specific needs were in a topic.”

Curriculum.

“And then, I always focus on standards. I don't deviate from the standards but you know you teach to the standards and the test. You know, unfortunately I don't feel like that's the right thing to do, but it is what you have to do in order to get them all in before those tests get here. And that's exactly what they're tested on and everything. You don't veer from that.”

Interview questions related to colleagues and teacher teams. These questions asked teachers to describe their interactions with colleagues about value-added scores and how their grade level or subject teams adapt or organize instruction in anticipation of value-added scores. All 19 participants described engaging in at least one interaction or practice with their colleagues regarding value-added evaluation. Two practices in particular were widely cited, departmentalization and leveling of classes.

Departmentalization at the elementary school level within the Community School District is an alternative arrangement to teaching a single, self-contained classroom. Instead, a teacher will instruct in one or two subject areas for two or more classes of students over the course of a daily or weekly schedule. Leveling is the practice of grouping students into self-contained classrooms or traveling homerooms (if the grade level is departmentalized) based on ability. The configurations vary, but the range of student abilities within the group of students is generally smaller than an intentionally mixed-ability class. Table 5.4 presents the categories of practices and interactions that were constructed from the initial codes.

Table 5.4. Participants' responses regarding interactions with colleagues

Question Specific Categories	Number of Teachers	Percentage of Teachers
Departmentalization	17	89%
Classes grouped by student ability	15	79%
Informal interactions	8	42%
Formal interactions	7	37%
Effect on relationships with colleagues	3	16%

Below are several quotations from teachers in response to these questions. Within each category, each quote is from a unique participant.

Departmentalization.

“I think that being focused [departmentalized] was a good thing and think that that idea came from TVAAS probably. In fact, I doubt they would've ever done that if it hadn't been for TVAAS.”

“Well it helps, you know. It helps the teacher to focus on just those one or two subject areas so that they can do better, plan lessons and stay more focused on those set of objectives as opposed to having so many standards that it's hard to keep up with and have all the kids master them.”

“We felt like that was to help our teachers become the master of a few subjects versus jacks of all trades. Because I would be more likely to be able to grow my students to their maximum if I was an expert give or take on a specific subject. Because I would be able to work with more tools and strategies if I was focusing on teaching reading and learning those strategies versus trying to be good at teaching reading, social studies, science, and math. That was the thought process.”

Classes grouped by student ability

“I still believe fully that leveling them in some, it doesn't need to be high, middle, low, but leveling them with sort of you know, students who are like-minded, helped because you know can meet them where they're at instead of trying to deal

with helping the low students and the high students are just bored and all that stuff. So that maximized the growth.”

“I've been in the classes where it was just not ability groups, and you had that high student and that low student and that low student immediately says I'll never get to where that kid is. It's demoralizing, and then you try to cooperative learn; you try all these, you know, techniques but in the end there's still no growth in that lower child. So I think ability grouping. I think our school does it to maximize growth.”

“But I think that I see the pendulum starting to swing back the other way a little bit. Maybe getting kids back into more heterogeneous groups so that there's more of a mix there. The idea of being maybe the ones that have some success can model more for the others who don't find as much success. There are advantages to both, I guess. But I see the pendulum swinging back towards more of mixing the kids up like we used to.”

Informal Interactions with Colleagues

“I have talked with our fourth grade teacher, [edited], and we actually, throughout the year she's my go-to and I'm her go-to, and we teach very similar, and we have ideas that are very similar, so when I was practicing for TCAPS, I asked what she

was doing and she's been my resource since she teaches science and social studies too, even though it's different standards and grade levels.”

“There wasn't much interaction at the beginning but when scores started coming out and certain people had higher scores than the other one, then there was a little more interaction like ‘what are you doing? what's your room looking like?’ So there was a little more interaction kind going on that way. Which is good.”

“Yes we talk about what websites and resources that we have seen to be helpful and you know getting them ready for the test.”

Formal Interactions with Colleagues

“We met. All of the science teachers and all the math teachers got together and talked about vertical planning. I thought that was really good.”

“We actually have PD [professional development] plans for the summer to map out our coursework for all of next year, to incorporate novel studies for the entire year and do this type of unit to teach everything next year. I definitely think that it will be the best avenue to grow our students because where we, where I teach at, most of our students do not have a lot of life experience.”

Negative Interactions with Colleagues

“I think it bred competition instead of collaboration in some teams.”

“Honestly, I had to take the stance of, once it was over and we had that beginning of the year meeting and talked about all of it, I would put it out of my head because I worked with another teacher who is constantly ‘TVAAS oh my goodness these kids are gonna ... da da da da ...’ constant. I was getting it from her, and I just had to leave that. It’s like I’m going to do what I need to do, and it’s going to be what it is.”

Interview questions related to principal interactions. These questions asked participants to reflect on their personal interactions with their principal regarding value-added evaluation. A total of 12 participants recalled interactions with their principals specifically about value-added scores. Some teachers reported more than one interaction. Table 5.5 shows three categories to describe the set of responses to these questions. Action-oriented interactions focused on steps teachers were given to improve student scores. Data-oriented interactions focused on explaining and reviewing student growth scores. Teachers who reported efficacy-oriented interactions focused on how the interaction impacted their beliefs of self-efficacy. Five teachers attributed efficacy-oriented interactions that included praise, encouragement and leadership opportunities from the principal to their favorable value-added scores. The other three teachers that

reported efficacy-oriented interactions with their principals experienced negative feedback regarding their efficacy.

Table 5.5. Participants' responses regarding interactions with principals

Question Specific Categories	Number of Teachers	Percentage of Teachers
Action-oriented interactions	3	16%
Data-oriented interactions	3	16%
Efficacy-oriented interactions	8	42%

Several quotations highlight the three categories, each from a unique participant.

Action-oriented interactions.

“She would come to us and she would say, ‘girls I've been looking at your scores’ and she would direct us, like she would tell us what we needed to do. So, I would have followed you know whatever, whatever she directed me to do ... I do specifically remember her coming to us every year, especially at the beginning of the year and saying ‘I want your focus to be on this, this year.’”

“The principal would suggest [teachers with poor growth scores] to go talk to somebody who had better scores. Go observe a teacher that had better scores ... I had [visitors] from other schools. Not in my, not in our school, but I had [the district math coach] bring teachers to my classroom ... and they were normally the schools [she] knew didn't do as well. She kinda made them, and she would take

them around certain teachers who were good at doing more stations and teach that way.”

Data-oriented interactions.

“So primarily when I came in, [the principal] sort of directed me in that this is where overall these kids are weak are this is the growth that they had achieved, I hope you will achieve at least one year's growth. You know that was always the goal, to see at least one year's growth.”

“The principal would meet with just a grade level and ... that was a stressful meeting too because we would talk about each teacher's language arts scores, math scores, science scores, and social studies scores and their value-added.”

Efficacy-oriented interactions.

“I was very lucky because all of my conversations were very positive because I was lucky to always have really good growth. So, the conversations we had were always very positive, which did help me to become somewhat of a leader in my grade level as well as my school.”

“Well I've got to be honest: the previous principal used to be very vindictive, you know this is what happened, this is what the TVAAS scores says about you, you need to do better. It was just kind of like the elephant in the room all the time.”

Interview questions related to school-wide practices. Participants were asked to reflect on their experiences with school wide practices that were used to improve value-added scores. Five categories were constructed from the initial codes. A practice that one participant referred to as “Data Day,” was cited by most of the teachers. “Data Day” is an annual meeting, usually held prior to the start of the school year, where subject and grade specific value-added scores were shared with the entire faculty. Another practice most participants had seen or experienced was the strategic placement of teachers in grade levels and subject areas with the intent to improve school TVAAS scores. Several participants noted that their school had a Data Response Team either currently or in the past to report the school’s value-added scores at “Data Day” or similar meetings.

Table 5.6 presents these categories.

Table 5.6. Participants’ responses regarding school-wide practices

Question Specific Categories	Number of Teachers	Percentage of Teachers
“Data Day”	14	74%
Strategic Teacher Placement	13	68%
Data Response Team (DRT)	5	26%
Additional Resources	3	16%
Testing Support	2	11%

The quotations below highlight the teachers’ responses to these questions. Each quote represents a unique participant.

“Data Day.”

“I had a principal who whose idea of looking at value-added scores was a data meeting with the entire staff. To look at scores and try to understand where the school needs to go, and I guess in a sense of this is where the school is filling out a matrix of quadrants of where we are achievement versus growth and where we need to work harder or smarter or what to work on this year.”

“She just brought [TVAAS data] right on out in front of everyone. So, I mean it was just up there. Every year right there at the beginning. Pretty or not pretty.”

“You didn't get an envelope. Everything was online, and all of the sudden everyone's meeting at the beginning of school looking at how your school did. And you know everybody's being held accountable whether you're first or second grade. You better, because third grade's struggling here and everyone's accountable, you know? It is everybody's job at your school, and it totally changed when the evaluation system came in and it came out public.”

Strategic teacher placement.

“I think a lot of it, like I was saying, and strategizing. So, you have an extremely strong ELA teacher you may not put them in third grade, but you may put them in fifth grade ... because third grade is the baseline.”

“Okay [participant’s name] has really high scores in third grade. Let’s move her to fourth grade to where we can have, so our school has, we can see growth. Then [participant’s name] gets moved to fifth grade. So that we can, when we are looking at that growth score, we have the highest growth in fifth grade. Not in third grade, not in fourth grade, but in fifth grade. I don't think that that's how it's intended, but I think that's how it was.”

“I was ready to move down and not be in a tested subject, and I also felt like my principal manipulated that to make me think I was ready to move down. Therefore, it was my idea and not hers even though she probably wanted me out of that position ... because she was not pleased with my reading test scores.”

Data response team.

“I'm part of the data team, the data response team ... We look at the data and take it back to our groups or our teams. We talk about when they roll out new curriculum. We talk about the dips and valleys that you're going to find in the curriculum, and it's important for me as a fifth-grade teacher to talk to the third grade and say ‘hey this is what I'm seeing. What are you doing in third grade that I can replicate in fifth grade?’ to maybe jog their memory. So that's a kind of strategy; I can't be everything that I need to be until you're everything that you are. I think Martin Luther King said that.”

“We do have a data team in math and reading, but I've not been involved in that, but they do meet throughout the year and discuss those benchmark tests in the end-of-the-year tests.”

Additional resources.

“I think the idea of trying to get TAs [teaching assistants] within classrooms was an intentional decision to try to maximize growth. The idea being that our teaching assistants were within your classroom to work directly with students.”

Testing support.

“Our school did something different [for TCAPs] than everyone else and try not to put any pressure on the kids. Because all throughout the day when the student came to their subject time slot so our schedules weren't really messed, I went to third grade and gave the test, the third-grade teacher went to fourth grade, and the fourth grade went to fifth grade ... I think it was to try to reduce test anxiety.”

Additional response categories not explicitly linked to interview questions. The participants were not asked specifically about problems with value-added methods (VAM), the Tennessee Value-Added Assessment System (TVAAS), or Tennessee Comprehensive Assessment Program (TCAP) standardized tests, but all 19 teachers

described at least one concern or issue with these topics. Table 5.7 presents the categories of issues the participants expressed in their interviews.

Table 5.7. Participants' responses regarding issues expressed

Question Non-Specific Categories	Number of Teachers	Percentage of Teachers
Questionable validity of VAM	13	68%
Negative effect on self-esteem	12	63%
Problems with application of policy	6	32%
Issues with TCAP	5	26%
Uncertainty and lack of understanding of VAM	4	21%
Negative effect on teaching	3	16%
Inter-school competition	2	11%

The quotations below highlight participants' responses not directly specific to any of the interview questions. Each quote represents a unique participant.

Questionable validity of VAM.

“Just a little bit I've studied, there are so many little nuances when you look at, like from tests of, like in high school. The test-to-test even from one grade to another, looking at coming from leveled classes to non-leveled classes, like when you try to compare growth from year-to-year, even kids within that same cohort, there are differences that happen within those years. That's just from what I've read, that it is hard to say that that's completely accurate or fair.”

“I think TVAAS is a corrupt, corrupt system ... TVAAS was implemented to evaluate, for lack of a better word, farming. In a controlled environment you could evaluate growth, you could evaluate everything as far as a plant ... So they have taken a TVAAS system and applied it to education where it has no business being in education because there are so many exterior factors that go into evaluation, and TVAAS is just so narrow-minded.”

“TCAPS, they focus on ‘every kid is supposed to make a year’s growth,’ okay? No matter what they walk in your classroom with, it should be a year's growth. My way, I feel that something would be more valid would be like if the test was administered the second week of school and that same test was administered at the end of the school year, but different questions and you could see if the kids grew. Because that's what you taught them, you know that's what you taught them. These are your standards and then did you do your job? Do they know more now than when you walked in your room? Not from what they knew from the previous year not what ... It's night and day from certain grade levels.”

“I never based anything about a kid solely on test scores. Because I know there's just kids that are great students and good, good kids but suck at taking a multiple-choice test. You know? Or have test anxiety, and I think such a big deal is made

about the stupid tests. I mean I've seen kids, I mean I've seen kids just shake. They were so nervous.”

Negative effect on teacher self-esteem

“You know, and hopefully there wasn't a lot of shaming. There have been years were teachers I think have been made to feel bad about it.”

“And at that point I felt like a failure of a teacher. And when I go into the following year after, you know, being told, shown these numbers, and everything feeling like I was less than, then it's really difficult to pick yourself back up and teach the best, the way you try every year. Like you feel like you're giving your all 100% all the time, and then those numbers are like the worst thing you can do to a teacher; it really kills their self-esteem, it kills how they feel about teaching in general, the children, it's just it's all around a negative. It's just negative all the way around.”

“I found it very upsetting. And the kids had made some growth; it just didn't necessarily show up in my TVAAS test scores. And that made me feel pretty worthless. I felt like I had to be the world's worst teacher.”

Problems with application of policy.

“I felt like she used them to determine our effectiveness in the classroom. And I feel like she also would after having seen our [TVAAS] scores then score us on our evaluation from her very similar to what our TVAAS was.”

Issues with Tennessee Comprehensive Assessment Program (TCAP)

“[TCAP testing measures] whether students can persevere on a task, whether they can work independently without any help from a teacher. And it measures their ability to solve problems independently and the ability to read a set of directions and understand what they're supposed to be doing, and it measures their ability to read, of course. I think that's about it. That's what it measures.”

“So that's just not, does not build good community within your classroom. And I think the stress is just horrible. I don't think that a third-grader, fourth-grader, fifth grader should be under that kind of stress to take a test that they know that it affects their scores and eventually will affect their teachers' work evaluation. I don't agree with it at all.”

Uncertainty and lack of understanding of VAM.

“I think that value-added component so secretive nobody actually knows the formula and what's put into it. Even though I know originally it was supposed to take a lot of the factors that can affect what goes on in your classroom and put them into some algorithm that generates some kind of a number ... some people

jumped on it at some particular point, and it kind of snowballed to a point where I think a lot of people used it but didn't understand it.”

Negative effect on teaching.

“It kind of goes back to what I said before: the love of teaching ... some of the creativity goes away because you're trying to hit those standards so well that it may not be quite as, you're maybe not as, I'm teaching them for the love of learning, it might be the love of get it done, you know?”

Inter-school competition.

“The way that it's been given in our county has not been helpful when you are rating 14 schools, one through 14 and how they scored. Teachers don't want to work collaboratively to grow, to improve, to get better because either you're number one in your county, and you want to stay there because that information is given to everybody, or if you're down in the bottom, then you don't want to ask for help because you don't want to talk about being at the bottom.”

Analysis of the Four Core Themes

As explained earlier in this chapter, grounded theory coding procedures were used to analyze the data in three stages: open, selective, and theoretical (Glaser, 1978; Urquhart, 2013) First, phenomena from the first several interview responses were assigned open codes. As interviews progressed, new codes were added, and existing

codes were consolidated through a constant comparative process (Glaser & Strauss, 1968) Through an immersive and re-iterative process, open codes were grouped into selective codes, or categories. Recurring categories were consolidated, and non-recurring categories were re-analyzed and adapted. See Appendix A for an illustration of the three stages of codes.

As the coding process continued, the data revealed that the participants' explicit understandings of value-added evaluation were diverse. At the same time, many commonalities among institutional and teacher practices surfaced across different schools within the district. Participants' widespread, unsolicited, and mostly negative perceptions of TVAAS and value-added evaluation emerged as well. An analysis of the phenomena, from both question-specific and non-specific responses across all participant interviews yields four core themes, illustrated in Table 6.1.

Table 6.1. Four core themes

Four Core Themes			
1 Teachers distrust value-added methods and evaluation.	2 Institutional practices establish norms in teacher evaluation.	3 Institutional practices related to value-added evaluations impact teacher self-esteem.	4 Teachers and teacher teams choose practices to improve student scores.

Theoretical codes were used to reorganize the selective categories into these four core themes incorporating the relevant phenomena for generating a theoretical model:

- Teachers distrust value-added methods and evaluation.

- Institutional practices establish norms in teacher evaluation.
- Institutional practices related to value-added evaluations impact teacher self-esteem.
- Teachers and teacher teams choose practices to improve student scores.

Theme One: Teachers Distrust Value-Added Methods and Evaluation.

Table 6.2. Core Theme One with categories

Core Theme One	Categories from Question Analysis
Teachers distrust value-added methods and evaluation.	Questionable validity of VAM
	Problems with application of policy
	Issues with TCAP
	Uncertainty and lack of understanding of VAM
	Negative effect on teaching
	Effect on relationships with colleagues
	Inter-school competition

All 19 participants shared concerns with value-added methods or evaluation policy, and most of the teachers questioned the validity of TVAAS. The data indicate distrust among teachers of VAM as a valid measure of teacher effectiveness or worth. The selection of quotations from participants illustrated this core theme. Each quote is from a unique teacher unless noted. Pseudonyms are intentionally not used to maintain confidentiality.

“I'm very honest with kids. I told them straight up; I told them all year long, ‘So these tests? I don't agree with these tests but we have to do it. You know if it were up to me we would have a better way to look at this but they don't let me make the rules.’”

“I mean you can't place a value on a teacher. I mean to me it's just an incredible waste of time to do value-added and attach it to a teacher.”

“You are touching a spot that I don't like ... Well if they were given a pretest and it was the same test then as the posttest, well then you're going to see growth. They take a totally different test, and I mean that's not really growth. It's a different test so I think it's stupid.”

“Sometimes I just feel like that number describes who I am as a teacher, but if, let's say, that year they just didn't perform well, or like we've had standards changes in math and reading. Like I feel like it's hard to know what the outcome's going to be.”

“I think when you're in a testing grade I think that it's always in the back of your mind, you are somewhat teaching to the test. You know? But I don't want to be that kind of a teacher; I don't want to be a teacher that teaches to the test. But I

want my kids to be successful so there's a really fine line because you really want them to be able to take that test to be successful, but you don't want to teach the test. So... it's a really bad ... situation for teachers in the upper grades.”

“You can see how that would build animosity. That's basically what TVAAS is doing in your schools you know. It's just making people angry with each other, or angry with the system, or angry at whoever's implementing the system.”

Theme Two: Institutional Practices Establish Norms in Teacher Evaluation

Table 6.3. Core Theme Two with categories

Core Theme Two	Categories from Question Analysis
Institutional practices establish norms in teacher evaluation.	Strategic Teacher Placement
	Data Day (Purpose)
	Data Response Team (DRT)
	Additional Resources
	Testing Support
	Efficacy-oriented interactions with principal
	Action-oriented interactions with principal
	Data-oriented interactions with principal

A majority of participants' accounts of institutional practices and interactions with their principals included the strategic removal/placement of teachers from tested grades to improve school scores, “public” displays of growth data to peers at annual faculty

meetings, and interactions with their principals regarding student growth scores. These practices were aimed at directly improving student growth scores through the strategic placement of teachers with histories of strong student growth on standardized tests or by motivating teachers to improve their performance and, in turn, improve their students' growth scores. The practices and interactions aimed at improving student growth scores legitimize valued-added performance measures. The selection of quotations from participants illustrates this core theme.

“It’s a stacked system...you want to stack your fourth and fifth grades...I know that principals tend to keep their, their very top teachers in those grades. And they move their teachers that are less effective to a grade level that is not tested. So you want like pretty good teachers in your third grade. Pretty good. You know, pretty consistent, pretty solid. And then you want really ‘on it’ teachers in your fourth grade. And again really ‘on it’ teachers in your fifth grade. And if somebody is less effective, then you move them to a subject or grade level or subject that does not [test].”

“Our math, that was our weakness. Third and fourth grade was really strong, but fifth-grade was struggling. So they moved fifth grade to third, that teacher, and then moved up the third and fourth grade teachers to fourth and fifth grade. So they like flipped and stayed in the content area, but they changed grade levels.”

“They display the scores, the TVAAS scores on a screen in the library during a meeting at the beginning of school year for everyone to see in each subject how they did and per grade level. And they talk about what they did wrong and how they can achieve that, and sometimes the teacher. That may be in mid-July when the scores are released. The teacher that taught that subject they did not have a score that the principal desired to have; sometimes [the teacher] got moved to a different grade level, so everyone would know why.”

“The principal and I did talk about growth a lot and what I could be doing to increase growth and what areas he was really interested in - statistics ... we talked about the statistics of all the numbers.”

Theme Three: Institutional Practices and Policy Impact Teacher Self-Esteem

Table 6.4. Core Theme Three with categories

Core Theme Three	Categories from Question Analysis
Institutional practices and policy impact teacher self-esteem.	“Data Day”
	Efficacy-oriented interactions with principal
	Negative opinions of TVAAS

All 19 participants reported on the impact of institutional practices and logics of value-added evaluation policy on self-esteem, including “Data Day” meetings, interactions with their principals about growth scores, and their personal value-added

scores. Six teachers shared narratives of “shaming” and “failure” as a result of school practices that made value-added scores “public” among the faculty. These six participants from three different schools recounted faculty meetings where teachers were described as “in tears,” and “upset.” Some participants reported negative emotions caused by interactions regarding TVAAS scores with principals, including “discouraging,” “embarrassing,” and “uncomfortable.” Four teachers reported positive experiences of self-esteem related to their TVAAS scores. Teachers also expressed negative feelings of stress and failure when reflecting on value-added evaluation. The selection of quotations from participants illustrates this core theme. Within each code, the quotations are from a unique teacher unless noted.

“Data Day.”

“There's one day where it's Data Day. And the whole faculty, K-5, the specials, everybody is there. And that's a great day if you did really well because everybody's going to look at what we did, and we look at all these graphs and charts, and there's a team that talks about the data and says how we did ... But if there's someone ... who did not do well, then it's really like ‘well, we didn't do well in fifth-grade math, and everything else was great, but fifth-grade math did not do well. So what was the reason?’ ... I mean I couldn't even, I couldn't even sit in there. I got to the part where they were like, ‘where did we fail?’ ... all the groups had to look at their paperwork and look at where we failed. And it was fifth-grade math, and it was me. And I couldn't sit in there. I was crying, and I had

to leave the room, and it was ridiculous and dramatic ... yeah, we had that meeting. It's an awful meeting.”

“It's humiliating because we're all sitting there talking about the scores, and [the principal] would say, ‘we need to improve.’ Well, people know who taught it, so you might as well say ‘[the participant] needs to improve her science score.’”

“One particular teacher comes to mind from the meeting. She just [got] up, having to leave in tears because she was so upset at what was being said in the meeting. Because at the time, the way that we were departmentalized at that time, it was that one teacher that taught that one specific subject at that grade level. Whereas now we have enough teachers teaching each subject in the grade level right now. But at the time, it was that, and she was visibly upset. I mean she's just having to get up and walk out of the meeting.”

Efficacy-oriented interactions with principal.

“I found it as an encouragement because this year I felt a little frustrated at times as I felt like [student scores] weren't where I wanted them to be. But having the conversations with [the principal] kind of told me that they are growing and so I just need to be patient. That what I'm doing is working and eventually they will get to where they need to be.”

“In our first PLC [professional learning community meeting] of the school year, [the principal] would give us all a sheet. It was a data meeting which had our class’ scores and rank. So, that was very discouraging. It was very uncomfortable. It was very embarrassing.”

“I had several years of really good value-added, and I had years where I didn't have good value-added. In the years that I obviously didn't have it, I was called into an office and ... it was never mentioned, ‘you have all these great scores when we come in and view you in your classroom, and you’re a great teacher’ you know? That's never discussed at that point. It's just this data, the numbers given.”

Negative impact of TVAAS on teacher self-esteem.

“I took my job very seriously and I think that's why I wanted to move out of, you know, being tested. Because I felt like I failed a lot of times. When I didn't see the results sometimes the scores made me feel like I did fail. I asked myself “What else can I do? What else can I do? What else can I do?”

“I always went in with that in mind. That there were going to be [TVAAS] scores there ... I didn't see a connection between what I did in the scores. I think it really didn't do anything but kind of stress me out.”

Theme Four: Teachers and Teacher Teams Choose Practices to Improve Student Scores.

Table 6.5. Core Theme Four with categories

Core Theme Four	Categories from Question Analysis
Teachers and teacher teams choose practices to improve student scores.	Test Preparation
	Planning
	Instruction
	Curriculum
	Departmentalization
	Classes grouped by student ability
	Informal interactions
	Formal interactions
	Individualized instruction

The data revealed a convergence of elementary school practices that teachers perceived to improve TVAAS scores or to avoid poor TVAAS scores. Eleven teachers attributed improvements in growth scores to departmentalizing. Participants expressed concerns about producing adequate growth in their students if they were required to teach all subject areas. Conversely, many of the participants expressed positive results and positive emotions when describing their experiences with departmentalizing. Fourteen teachers expressed a preference for teaching one or two subjects over four or more subjects in self-contained classrooms. Several participants said they were better able to grow their students and better able to help students master the content with more “focus”

and fewer standards to teach. Participants described themselves as experts and as having expertise in their subject areas. In response to the volume of standards and a perceived lack of time and resources, several teachers expressed “relief” from the responsibility of teaching standards in all subject areas. Teachers described their preference for teaching a single or pair of subjects fitting their perceived strengths. Some teachers reported a desire to leave tested grades or subjects where they have not performed well. Nine of the 19 participants reported that they had relocated to teaching positions that do not include individual growth scores as a component in evaluation.

Participants also reported collaborating with colleagues both formally and informally. Team practices also included preparing students for standardized tests and adapting curriculum and instruction. Some teachers reported that they kept the test “in mind” with some ambivalence when planning and teaching. Eleven of the participants reported that they allocated significant instructional time to prepare students for annual standardized tests. Teachers also reported planning curriculum and instruction with standardized test content, format, and language in mind. Teachers described how they used class time to administer practice tests with similar format and conditions to improve students’ performance on the actual tests. While some teachers reported that leveling students by ability was effective for improving growth scores, others were dissatisfied with levelled classrooms. Several teachers reported using individualized instructional approaches to improve growth scores by providing extra help when needed and strategically focusing teaching resources on “bubble kids,” or students who had the most potential to show growth on standardized tests.

Test preparation, planning, curriculum, and instruction.

“You always have that test in mind. If we thought that something might be good for the students but it would be detrimental to testing, we wouldn't have done that ... I can't think of an example of that right off the top of my head. But it would've been a definite 'no.' Got to show growth on the test always, that's always there.”

“Instead of maybe spending more time with hands-on area and perimeter things, we were more or less trying like to figure out what are all the different ways of figuring out a [test] question.”

“For example, 4th grade tests. I know the largest percentage is going to be in fractions. And so, knowing that they can only use a basic calculator this year, unlike last year, I had to spend a lot of time with one of my classes making sure that they had the solid foundation in fractions so that they might perform better on the test. Before I could move on to things they were more interested in like geometry, I spent the least amount of time on geometry because it's something that they've been exposed to for the longest. So, yeah is just a difference in the students that I have and knowing really what your standards are and knowing the test is driving instruction in the aspect of having to spend more time on fractions because it's a larger portion of the test.”

“I mean the vocabulary comes straight out of my standards. We focus on that through the whole year. So, I think that I know what they are tested on, so we keep building on that throughout the year. So it's constantly reinforced, so when they do take the test, they're used to those terms.”

“I know that coming from [former school] where we were self-contained, our scores weren't as good as a fourth-grade team or as a fifth-grade team as they have been since we were departmentalized.”

Departmentalization.

“I just didn't feel like I had, I wasn't impacting the students as much as I am now. I'm just teaching math, and it's my favorite subject, and I've got a degree in math, so that helps with my teaching. Math is my favorite.”

“Because there is so much content to cover with the new standards, I believe you have to spend so much time planning ... by only having to look at two subject matters, I have more kids to look at, but I can really work on the science of teaching math, for instance.”

“But I mean, I remember when we first departmentalized, that we were so excited. We were really stoked. You know, it was like ‘Oh! Finally, I can focus on one area, one thing.’ You know? And I was like excited again ... I think we all ended

up [in the subject area] where we should be. There again our test scores probably reflected that was our strength and our area of interest, and we are excited about it, and thus, I think we teach it better. It's more engaging so I think that probably drives test scores up versus the test score dictates where we are.”

“I know for me as an upper grade level teacher I would not be comfortable trying to teach all four core subjects to my students and I would be very nervous about what their growth would be because I am not as learned in all the subjects, if that makes sense. And I'm not as comfortable, and I feel like for me, when I look at growth and learning, I need to be comfortable or somewhat knowledgeable in the subjects so that I can handle their questions. Or at least feel like I've done everything I know how to do, so let me go ask somebody else. And if I had to do that with all of my subjects, I don't know that my students would get the best of me, so to speak, because I would still be trying to learn all four areas, content areas, instead of just focusing on two.”

Individualized instruction.

“There were times where I kept students in for like 15 minutes of recess, and we would work together on specific things. I offered tutoring two days a week in the morning for an extra 30 minutes. So the kids would come in. I'd handpick the students that I wanted to come in...looking at the TVAAS scores... Those were my bubble kids, that I could get a little more growth out of and that needed a little

bit more growth. And I also offered two days a week, afternoon tutoring sessions and handpicked those students as well ... I just really focused on that group of students to try to get a little growth where I could.”

“[The principal] would tell us to focus on the ones that we could move. She's like, ‘if you think the student is close enough to the line, that you can move them over the line, that's who we need to be focusing on.’ Like in our benchmark testing ... the student was in the yellow (near proficient), but they were close to the green (proficient); go ahead and try to push them up on over into the green. So put a lot of focus on that student. Whereas if you had a student that was in the yellow, but their scores were low enough that they were almost in the red (well below proficient), then at that point in time, our focus would not be on that student. We would still be teaching them, doing our best with them, but we’ll be pushing the ones that are closer to the line ... we based it off of DEA (Discovery Ed Assessment) scores because those were scores that we could look at. But it was the intent ... to move those medium students up higher. So we based it off of the DEA, but the intent was to get them to show growth.

“There's more review if they don't perform well enough or sometimes I have an assistant pull out students to catch them up if somehow they missed getting that concept because I know that's going to affect ... because I want them to learn, yes, but also because I want them to have a positive value-added, for us to have a

positive value-added. I mean you want every student to grow that two points or more, but you want as an average the whole class, you know, to do well.”

“So by individualizing instruction for the small group, we felt like we’d be able to maximize that growth potential more.”

“This year I would say that teaching-wise there has definitely been thinking about how we’re going to address all of the standards so that all of my students understand. So style of teaching ... I still teach the way I did even in kindergarten, I still teach their centers and through rotations, but it has to be more focused on the individual student and where they are and the individual standard to make sure they're hitting all of the standards to be prepared for the test.”

Development of Theoretical Model

Table 6.1. Four core themes

Four Core Themes			
1 Teachers distrust value-added methods and evaluation.	2 Institutional practices establish norms in teacher evaluation.	3 Institutional practices related to value-added evaluations impact teacher self-esteem.	4 Teachers and teacher teams choose practices to improve student scores.

Using grounded theory coding procedures, the data were analyzed in three recursive stages: open, selective, and theoretical (Glaser, 1978; Urquhart, 2013). First, phenomena from the first several interview responses were assigned open codes. As interviews progressed, new codes were added, and existing codes were consolidated through a constant comparative process (Glaser & Strauss, 1968). Through an immersive and reiterative process, open codes were grouped into selective codes, or categories. Recurring categories were consolidated, and non-recurring categories were re-analyzed and adapted. Drawing from Corbin and Strauss' (2008) framework for analyzing data for process, the categories were analyzed to develop a theoretical model that explains the participants' adaptation to value-added evaluation. This process model represented the sequence of actions, interactions, and emotions that changed in response to combinations of contexts, conditions, events, and interactions based on the researcher's interpretation of the data (Corbin & Strauss, 2008).

The data revealed that participants' explicit understandings of the intent of value-added evaluation policy were divergent. Surprisingly, several institutional, team level, and classroom practices that were adopted in response to value-added evaluation were convergent across schools, despite the absence of a district policy mandating the practices. Participants' widespread, unsolicited, and mostly negative perceptions of TVAAS and value-added evaluation surfaced as well. An analysis of the phenomena, from both solicited and unsolicited responses across all participant interviews, yielded four core themes. Theoretical codes were used to reorganize the selective categories into

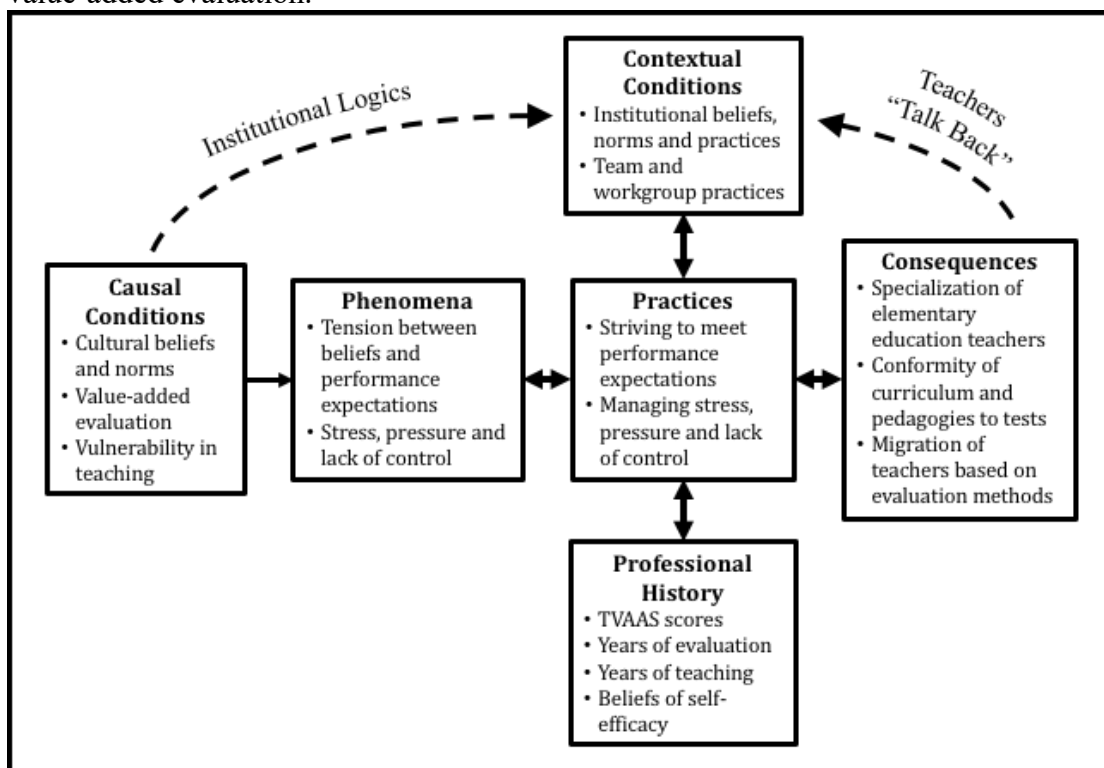
these four core themes that could incorporate the relevant phenomena for generating a theoretical model:

1. Teachers distrust value-added methods and evaluation.
2. Institutional practices establish norms in teacher evaluation.
3. Institutional practices related to value-added evaluations impact teacher self-esteem.
4. Teachers and teacher teams choose practices to improve student scores.

Development of the Theoretical Model

Four core themes were constructed from the study's findings. Integrated back into a whole, these four core themes yielded a theoretical model that explained how elementary education teachers have responded to Tennessee's value-added evaluation policy through decisions about their practice. This theoretical model is presented in Figure 4.

Figure 4. Theoretical model to explain how teachers choose practices in response to value-added evaluation.



Causal conditions of phenomena related to teachers’ response to value-added evaluation. Three types of causal conditions surfaced from the literature review, which led to phenomenological experiences related to value-added evaluation. These causal conditions are:

- cultural beliefs and norms
- value-added evaluation
- vulnerability in teaching

Cultural beliefs and norms of standardization of meritocracy, free market competition, and results-based accountability in the public sector intersected to create a policy

environment for value-added evaluation in education. These ideologies assumed a direct causal relationship between teacher quality and standardized measures of student learning. Macro-level discourse and evaluation policies that link student growth scores using standardized measures with high stakes consequences or incentives for teachers have legitimized standardized test scores as fair and efficient measures of student learning and teacher quality. The second causal condition consisted of personal experiences with value-added evaluation. To be selected for the study, teachers were either currently being evaluated by TVAAS or had been previously evaluated by TVAAS in their professional history. The third causal condition was the structural condition of vulnerability in teaching. The data suggested widespread distrust among the participants toward value-added models as valid measures of teacher effectiveness or worth. Teachers did not agree with the underlying assumption of TVAAS that they have full control over their students' growth scores. As a result, value-added evaluation policies could exacerbate teacher perceptions of professional vulnerability (van Veen et al., 2005). This finding is consistent with other research showing that results-based accountability systems in education reform have evoked high levels of uncertainty, instability, and vulnerability among teachers (Ball, 2003). Teachers gave a variety of reasons why they believe TVAAS was not a useful measure for teacher evaluation. Some participants expressed that standardized tests were not valid instruments for measuring what students have learned or teacher effectiveness.

Phenomena reflecting cultural norms, value-added evaluation, and conditions of vulnerability. Causal conditions—cultural beliefs and norms, value-added

evaluation, and vulnerability in teaching—contributed to two categories of subjective phenomena as reported by participants:

- tension between beliefs and performance expectations
- feelings of pressure, stress, and lack of control

Teachers' experiences of vulnerability caused them to doubt the validity of value-added methods that claimed to control for all factors in a student's learning, to isolate a teacher's contribution to student learning, and to measure their value-added accurately against other teachers. When faced with an evaluation policy that measures efficacy based on standardized meritocracy and market-based ideologies, teachers experienced a tension between structural pressures to produce and their perceived lack of control over the production outcomes. Teachers also experienced stress, pressure, and lack of control as a result of the causal conditions. Teachers shared negative emotions when reflecting on value-added evaluation, including "embarrassing" situations, low school morale, and frustration. Participants also experienced pressure in "Data Day" meetings because "scores are broadcasted."

Contextual conditions. Teachers' agency in developing strategies for choosing practice was mediated by contextual conditions. Strategies for choosing practices were developed in response to the two subjective phenomena. These strategies were influenced by contextual conditions including:

- institutional norms and practices
- team and workgroup practices

Teachers expressed emotions of stress, pressure, and failure when reflecting on their experiences with TVAAS scores that did not meet institutional expectations. Most of the participants shared narratives of “shaming” or “failure” as a result of school practices that made value-added scores “public” among the faculty. Multiple teachers from different schools recounted faculty meetings where teachers were “in tears,” and “upset” because their TVAAS scores had not shown a year’s worth of growth. Some participants reported other emotions caused by interactions with principals, including “discouraging,” “embarrassing,” and “uncomfortable.” The principals and teams shape the contextual conditions for teachers. These conditions influence how teachers choose their practices and what practices become integrated into the contextual conditions, creating a recursive cycle.

Teachers described several institutional practices and interactions with principals, including the strategic placement and removal of teachers from tested grades to improve school scores and “public” displays of growth data to peers at annual faculty meetings. In addition, teachers recounted conversations with their principals who encouraged them or sought to motivate them to improve their students’ growth scores. These institutional practices were reported across several schools in the study. Teachers also described a wide range of team and workgroup practices including departmentalization, leveling of students by ability, formal vertical planning across grade levels, and sharing of teaching and testing strategies through informal interactions.

Professional history. In addition to contextual conditions, elements of teachers' professional histories also mediated their agency in choosing strategies. These elements included:

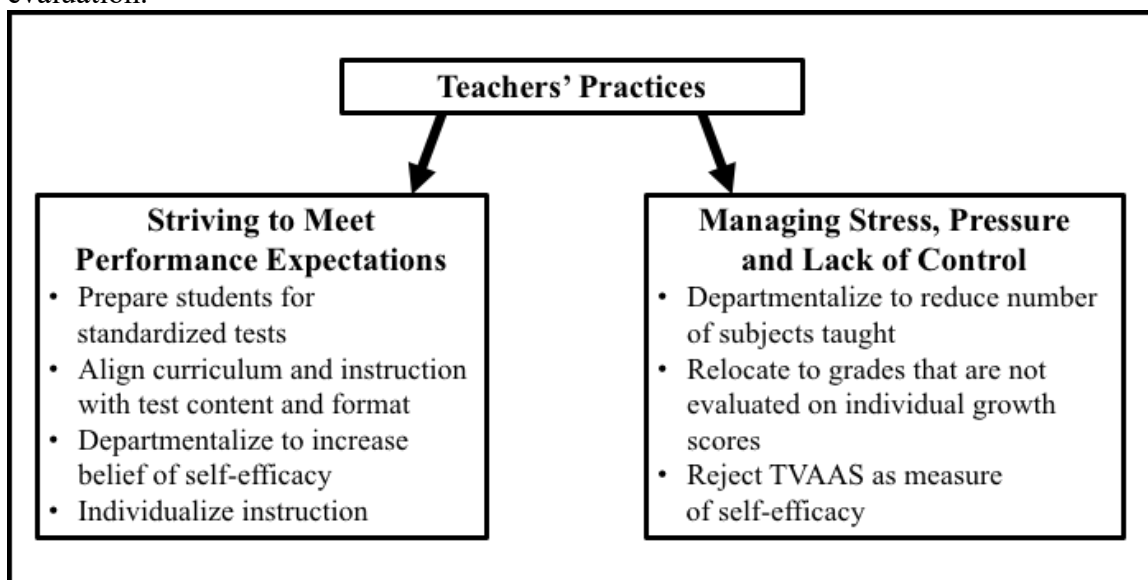
- TVAAS scores
- Years of evaluation
- Years of teaching
- Beliefs of self-efficacy

Strategies for choosing practices. Two phenomena, mediated by the contextual conditions and teachers' professional histories, led to the development of two categories of strategies for choosing practices:

- striving to meet performance expectations and
- managing pressure, stress, and a lack of control.

The strategies and practices are presented in Figure 5.

Figure 5. Teachers' strategies for choosing adaptive practices in response to value-added evaluation.



Striving to meet performance expectations. The data revealed that many of the teachers in the study have adopted several strategies aimed at meeting performance expectations including:

- developing students' test taking skills
- aligning curricula and instruction with test content and format
- increasing efficacy through departmentalization
- individualizing instruction

Given teachers' widespread experiences with institutional practices that triggered feelings of shame and failure for teachers who did not meet institutional expectations for annual growth, many teachers chose practices they believed would help them reach performance expectations.

Developing students' test taking skills. Teachers adopted several practices aimed at improving their students' performance on standardized tests, including efforts to prepare students for standardized test format and conditions. Teachers' practices to prepare students for standardized tests included dedicating instructional time to students practicing with state produced tests, creating their own questions to mimic test items, and simulating test conditions for their students.

Aligning curricula and instruction with test content and format. Teachers also aligned curriculum and classroom instruction to include vocabulary expected on standardized tests, focused skill practice, and pedagogies that they believe will translate into improved test scores. Some teachers reported adapting the pacing of instruction to maximize test scores as well.

Increasing self-efficacy beliefs through departmentalization. Many of the participants expressed an increased belief of self-efficacy when sharing their experiences with departmentalizing. Teachers described themselves as an "expert," "knowledgeable," and "comfortable" when they taught fewer subjects. Most of the participants reported how teams of teachers have adapted teaching configurations to reduce the number of subjects they teach in response to the high volume of standards and a perceived lack of time and resources to master them. Teachers felt "relief" from the responsibility of teaching standards in all subject areas. Several participants said that with more "focus" and fewer standards to teach, they were better able to grow their students and to help students master the content. While participants were not asked if they preferred

departmentalization, nearly all of the teachers' expressed a preference for teaching one or two subjects over four or more subjects in a self-contained classroom.

Individualizing instruction. Teachers indicated that they individualized their instruction to meet students' specific learning needs in order to maximize growth. Some teachers enlisted teaching assistants to provide extra instruction to help students "catch up" when they did not master standards with the rest of the class. Several teachers reported using individualized instructional approaches to improve growth scores by strategically focusing more teaching resources on "bubble kids," or students who had the most potential to show growth on standardized tests, and fewer resources for students who appeared less likely to progress to a higher proficiency category.

Managing stress, pressure and lack of control. Several participants shared that they had changed teaching positions to grades where their students' TVAAS scores were not used to calculate their evaluations. These changes were initiated by the teacher's request or his or her principal. Some teachers reduced the number of subjects they taught by departmentalizing. Other teachers reported leaving tested grades or subjects where they did not perform well or experienced unwanted "stress" and "pressure." These teachers found that changes in teaching assignments reduced experiences of stress and pressure about value-added scores. Some teachers rejected TVAAS as a measure of efficacy and intentionally did not change their teaching practices. Other teachers were relocated by their principals to untested positions. While all 19 participants had been evaluated using individual growth scores in the past, only eight of the 19 participants reported TVAAS as part of their evaluations at the time of their interview. It was also

worth noting that nine of the 19 participants reported relocating to teaching positions without individual growth scores at some point in their professional history.

Consequences for strategies of choosing practices. The strategies used by some participants brought about consequences in elementary education practice, including a specialization of teachers to a limited number of subjects areas, a shift to curricula and pedagogies that aligned with test content, opportunity costs of instructional time dedicated to test preparation, and value-driven performances by teachers replacing values-driven practices (Ball, 2003). Some of these strategies, including departmentalizing, relocating to another grade, and rejecting TVAAS as a measure of self-efficacy, mitigated teachers' phenomenological experiences with failure, shame, pressure, and stress. In addition, some teachers chose strategies they believed had improved their self-efficacy to meet student growth expectations, including preparing students for tests, departmentalizing to build expertise, and individualizing instruction to improve student mastery of skills. These strategies reflect how teachers exercised agency in policy implementation through their choices of practice within the institutional constraints and pressure of evaluation policy.

While their strategies for choosing practices may have increased teachers' sense of agency and reduced stress and pressure when teachers felt they failed, teachers continued to distrust value-added evaluation and to struggle with its perceived impact on teaching and their students. The data suggested little confidence in value-added models as valid measures of teacher effectiveness or worth. Teachers gave a variety of reasons why they believed TVAAS was not a useful measure for teacher evaluation. Some participants

expressed that standardized tests were not valid instruments for measuring what students have learned or teacher effectiveness. One teacher referred to the tests as “an incredible waste of time.” Others shared their frustration with having to “teach the test.” Additionally, most of the teachers in the study suggested teacher improvement as the best use of evaluations, but they did not find their value-added scores to be helpful for improving their own teaching.

Summary

The four core themes from the interview data included:

1. Teachers distrusted value-added methods and evaluation.
2. Institutional practices established norms in teacher evaluation.
3. Institutional practices related to value-added evaluations impacted teacher self-esteem.
4. Teachers and teacher teams chose practices to improve student scores.

The four core themes were analyzed to generate a theoretical model to answer the research question: “What theoretical model can explain how teachers have adapted their practices in response to value-added evaluation?”

The theoretical model constructed from the four themes to answer the central question proposes that:

Three causal conditions (1) cultural beliefs and norms, (2) value-added evaluation, and (3) vulnerability in teaching led to subjective phenomena of (1) tension between perceptions of agency and performance expectations and (2) feelings of stress, pressure

and lack of control. Strategies for choosing practices in response to these phenomena were (1) striving to meet performance expectations and (2) managing feelings of stress, pressure, and lack of control. Teachers' agency in choosing strategies was mediated by their contextual conditions including (1) institutional beliefs, norms and practices and (2) team and workgroup practices. The development of strategies was also influenced by a teacher's professional history which included (1) TVAAS scores, (2) years of evaluation, (3) years of teaching, and (4) beliefs of self-efficacy. Professional histories and contextual conditions shaped two parallel strategies for adapting to value-added evaluation: (1) striving to meet performance expectations and (2) managing stress, pressure, and lack of control. Strategies for striving to meet performance expectations included (1) preparing students for standardized tests, (2) aligning curricula and instruction to test content and format, (3) departmentalizing grades by subject area to increase belief of self-efficacy, and (4) individualizing instruction. Strategies for managing feelings of stress, pressure, and lack of control included (1) departmentalizing grades by subject area to reduce the number of subjects taught, (2) relocating to grades or subjects where personal TVAAS scores were not used for evaluation, and (3) rejecting TVAAS as a valid measure of self-efficacy. The consequences of these strategies included (1) specialization of elementary education teachers, (2) a shift to curricula and pedagogies that aligned with test content, (3) the opportunity costs of instructional time dedicated to test preparation, and (4) value-driven performances by teachers replacing value-driven practices. These strategies mitigated teachers' phenomenological experiences with failure, shame, pressure, and stress and empowered teachers to exercise

agency through improved beliefs of self-efficacy. However, teachers continued to distrust value-added evaluation and struggled with its perceived impact on their teaching and their students.

This chapter included summaries of how the data were collected, the demographics of the participants, the methodology followed for analyzing the data, the analysis of the responses to interview questions, the analysis of the four core themes, and the synthesis of the theoretical model. Chapter 5 explores the meaning of the study's findings in response to the central research question to address the implications of the findings in the context of teachers' lived experiences, elementary schools, and the broader educational field. Theoretical and conceptual frameworks introduced in the literature review will be discussed in relation to the findings and research question of this study. Implications for practice, policy, and theory will be presented as well as suggestions for future research. The chapter concludes with a personal reflection on how the learning journey of this study has transformed the researcher personally and professionally.

CHAPTER V

CONCLUSIONS

The purpose of this study is to develop a model to explain teachers' responses to value-added evaluation policies (VAEP) through their reflections and narratives as empirical research to understand policy dynamics and teachers' roles in the implementation process. By gathering teachers' accounts of their beliefs, perceptions, experiences, and practices related to VAEP, this constructivist grounded theory study offers a theoretical model that explains the process teachers experience in choosing their practices in response to a state-mandated, value-added teacher evaluation policy.

The data collection for the study is guided by the following research question:

- What theoretical model can explain how teachers have adapted their practices in response to value-added evaluation?

From this central question, two empirical questions followed:

- What are the lived experiences of teachers who have been evaluated by Tennessee Value-Added Assessment System?
- What practices do teachers choose in response to Tennessee Value-Added Assessment System evaluation policy?

Review of Methods

The results presented in this chapter are drawn from single, semi-structured interviews with a theoretical sample of 19 elementary education teachers from the

Community School District, a K-12 school district in Tennessee, a state that implemented a value-added component to teacher evaluations in 2011. Interviews are arranged once participants provide informed consent. All participants are interviewed by phone for consistency in data collection and to reduce potential bias in participation based on logistical challenges of meeting face-to-face. Through constant comparative analysis, additional participants were recruited using these protocols to reach theoretical saturation. Participants were interviewed once within a three-month period between April and June of 2017 to improve dependability of responses (Shenton, 2004). This time period began after state standardized testing and ended before scores were reported. All phone interviews were audio recorded and transcribed for analysis.

Using grounded theory coding procedures, the data were analyzed in three recursive stages: open, selective, and theoretical (Glaser, 1978; Urquhart, 2013). Phenomena from the first several interview responses are assigned open codes. As interviews progress, new codes were added and existing codes were consolidated through a constant comparative process (Glaser & Strauss, 1968). Through an immersive and reiterative process, open codes were grouped into selective codes, or categories. Recurring categories were consolidated, and non-recurring categories were re-analyzed and adapted. Through focused coding, (an update of axial coding) codes that can account for both convergences and divergences in the data were identified (Charmaz, 2014; Saldaña, 2012). These most salient codes were clustered together into emic, situated categories to produce four core themes that represent the relationships between the codes (Carley, 1993; Maxwell, 1996; Saldaña, 2012). Drawing from Corbin and Strauss' (2008)

framework for analyzing data for process, the categories were analyzed to develop a theoretical model that explains the participants' adaptation to value-added evaluation. This process model represents the sequence of actions, interactions, and emotions that change in response to combinations of contexts, conditions, events, and interactions based on the researcher's interpretation of the data (Corbin & Strauss, 2008b).

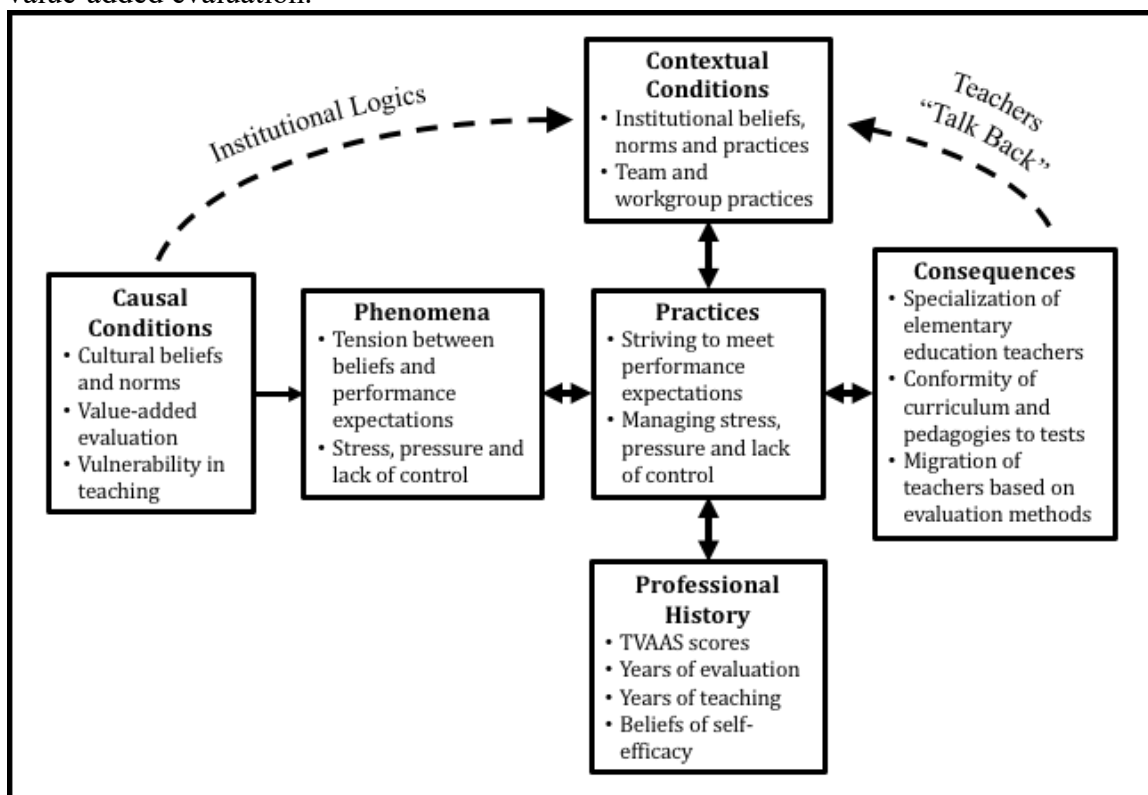
Discussion of Findings

Table 6.1. Four core themes

Four Core Themes			
1	2	3	4
Teachers distrust value-added methods and evaluation.	Institutional practices establish norms in teacher evaluation.	Institutional practices related to value-added evaluations impact teacher self-esteem.	Teachers and teacher teams choose practices to improve student scores.

The findings in Chapter IV described four core themes that were generated from the open and selective codes. These core themes are presented in Table 8. The four core themes are integrated to construct a theoretical model to explain teachers' choices in practice in response to value-added evaluation. Drawing from Corbin and Strauss' (2008) process framework for grounded theory, a theoretical model is generated from these core themes to respond to the central research question. The theoretical model is presented in Figure 4.

Figure 4. Theoretical model to explain how teachers choose practices in response to value-added evaluation.



The theoretical model constructed from the four themes to answer the central question proposes that there are:

three causal conditions:

- cultural beliefs and norms,
- value-added evaluation, and
- vulnerability in teaching, which lead to

two subjective phenomenological experiences:

- tension between beliefs and performance expectations and
- feelings of stress, pressure and lack of control.

In addition, teachers' agency in choosing practices is mediated by their contextual conditions including:

- institutional beliefs, norms and practices
- team and workgroup practices

Furthermore, the development of practices is also influenced by teachers' professional histories, which include:

- TVAAS scores,
- years of evaluation,
- years of teaching, and
- beliefs of self-efficacy

Therefore, subjective phenomenological experiences, contextual conditions and teachers' professional histories interact and shape two parallel categories of practice for adapting to value-added evaluation:

- striving to meet performance expectations, and
- managing stress, pressure and lack of control.

Practices chosen for striving to meet performance expectations include:

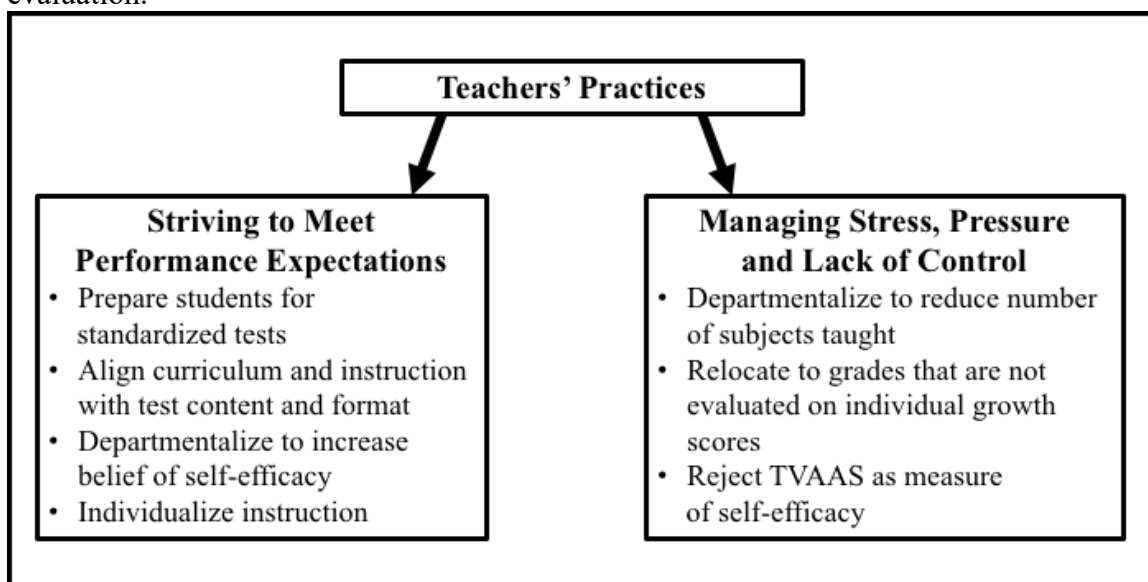
- preparing students for standardized tests,
- aligning curricula and instruction to test content and format,
- departmentalizing grades by subject area to increase belief of self-efficacy,
and
- individualizing instruction.

Practices for managing feelings of stress, pressure and lack of control include:

- departmentalizing grades by subject area to reduce the number of subjects taught,
- relocating to grades or subjects where personal TVAAS scores are not used for evaluation and
- rejecting TVAAS as a valid measure of self-efficacy

Teachers' practices are presented in Figure 5.

Figure 5. Teachers' strategies for choosing adaptive practices in response to value-added evaluation.



This chapter provides a response to the empirical research questions and situates the theoretical model within grounded theory reliability criteria, within the conceptual framework presented in the literature review, and within value-added evaluation policy literature discussed in Chapter II. Discussions of the implications for practice and policy as well as recommendations for further research are also included. The chapter concludes with the researcher's reflection and closing remarks.

Responses to Empirical Research Question #1

- What are the lived experiences of teachers who have been evaluated by Tennessee Value-Added Assessment System?

Perceptions of value-added evaluation policy. Participants are asked what they believe the intent of the policy to be. While their responses provide explicit understandings of policy (or lack thereof), the participants' narratives of institutional practices relate to value-added evaluation and their phenomenological experiences provide rich data about their understanding of how the policy has been implemented in their schools. Direct responses about policy intent vary from language nearly verbatim from the policy, to more natural language, to admitting a lack of understanding of its intent:

“It's a very basic level to ensure growth of students or to monitor growth of students. And then to attach that to, you know, evaluating whether a teacher is more or less effective.”

“... it takes your students where they started and it shows what you actually have accomplished, what they have accomplished through what you've taught.”

“I guess it's just like you know once again it's some kind of an equation that they put in place that this is you know, when we plug these numbers and this is where it comes about, you know, this is your number this and that. I really don't

understand how the formula that they have, how that all works. I just know when our union had like sat down and had meetings before, they had someone explaining why it doesn't work, and I don't really understand that component, I know it's not right, but I don't understand like why that equation or that formula is the formula that it is.”

Most responses from teachers who could articulate intent of policy expressed three categories: to improve teachers, to improve student growth, and to hold teachers accountable:

“I think that they want teachers to look at the growth to see what practices worked and didn't work in order to help those students perform well on the test. And I think that's what it is it is for you to look back at your data and see what you can do differently to improve on.”

“I guess to see if you are being an effective teacher, because your kids will grow.”

“I believe because I'm sure in their brain it's a way to hold teachers accountable and push them to make sure that they're working harder than the teacher before.”

Most participants also volunteered opinions and experiences of value-added policy that were not related to its intent. Most teachers expressed distrust of the validity

of value-added evaluation methods and its usefulness. Others expressed concern with its effect on relationships between teachers and schools:

“I feel like, you know that when [principals] come into my classroom, that's a better, that is a much better picture of how I teach, of how my kids are learning than a TCAP score ever would be ... so, I don't like testing.

“You know I think teachers get a high score one year and the next year, a low score. You know and the next or they go back up to high score. So I mean you look for patterns over time, but I mean a lot of what you do is things that aren't in your control also.”

“The problem is I guess getting evaluated with the type of evaluation we have, it doesn't include like the type of economic status of students and their home life and where they come from. Like I think that's what it's lacking because that's the problem with using it.”

“I don't think is appropriate because teachers have to choose what scores they want to use without even knowing what the scores are. It's ridiculous, for your principal in February [to] say ‘what you want to use for your ... value-added? Do you want these value-added from math and science?’ And how can you say what

you want to use it for because you don't even know how they did? Yeah I think it's ridiculous.”

“I think this really tying these scores to teacher evaluations I feel like greatly impacted school morale. In a negative manner. Instead of it being used to positively impact students.”

“I think you can definitely create [a competitive situation]. I remember asking [school colleague] how I compared to [team colleague]. And she's like, “you two teach very similar. So, I'm like, then what's the difference?” And then when she moved to fifth grade she found out it's a harder curriculum.”

Phenomenological experiences of teachers related to value-added evaluation policy. Questions about institutional practices related to value-added evaluation evoked many narratives of phenomenological experiences of failure, shame, pressure, and stress. Several teachers shared positive experiences with favorable TVAAS scores. Teachers also described changes in their experiences of teaching since TVAAS evaluations:

“There have been years where people I think have felt ashamed.”

“I found it very upsetting. And the kids had made some growth it just didn't necessarily show up in my TVAAS test scores. And that made me feel pretty worthless. I felt like I had to be the world's worst teacher.”

“If my kids really [struggle] with something, I'm left with a dilemma. Well do I keep going with this until they get it? Because they just need more practice to get. Do I stop and go on so that we can get through all of the standards before testing? And I think that's a very difficult and unfair place to put teachers. So, I think it has that negative effect.”

“And the stress that was put on principals. And you know it trickles down and then we put the pressure on the kids. You know everybody has that pressure all over a score. All over when that report comes out.”

“I understand the need for us to have testing and standards, but I do feel like it takes a little bit away from the joy of teaching...I have definitely seen in many of my colleagues and myself how you will try harder if you know that you are accountable.”

“In the past years I see the art being taken out of it. That has frustrated me a lot...I think it was it was coming from our standardized testing scores. Because that was mentioned we've got to get the scores up. We've got to get the scores up by this

literacy leader and that was her feeling was that we weren't teaching the skills. So we were broken down into teach the skills.”

By considering teachers' understandings of policy intent along with their accounts of phenomenological experiences related to value-added related institutional practices, the data suggest that teachers perceive that value-added policy is unable to accomplish its goals of improving teachers, to grow student achievement, or to hold teachers accountable. Teachers' self-esteem is positively and negatively affected by the institutional practices related to value-added evaluation, depending on their actual TVAAS scores. Many teachers experience feelings of failure, shame, pressure, stress, and dissatisfaction in teaching when low scores are attributed to their name, especially when publicly displayed and discussed among their peers. Conversely, some teachers believe that their high TVAAS scores lead to principals' offers of leadership and mentorship opportunities, praise, and more freedom in the classroom.

Responses to Empirical Question #2

- What practices do teachers choose in response to Tennessee Value-Added Assessment System evaluation policy?

Teachers have developed two strategies for choosing practice:

- striving to meet performance expectations
- managing stress, pressure and lack of control

Teachers strive to gain control over student outcomes by adopting several practices they believe will improve their students' standardized test scores. These

strategies include developing their students' test taking skills, aligning curricula and instruction with test format and content, increasing efficacy through departmentalization, and individualizing skill-based instruction:

“We are always practicing test taking strategies in everything that we did but the actual practicing would come closer to the test.”

“I would do some practice tests closer to spring, closer to the time when they were to take the test and not so much content-wise. Well, obviously there's content--I did a lot for just practice taking that style of the test ... since that's how we test, I would certainly do some practice tests sometimes; I would do them at like at the end of the first semester we would do one. And then we would do it again a couple of weeks before they took the actual test.”

“I think when you're in a testing grade I think that it's always in the back of your mind, you are somewhat teaching to the test. You know?”

“Oh, I love having the different departments. I think it makes each teacher a [specialist] in their subject ... we kind of go directly with our subjects because that's where strengths are ... altogether I love how fifth grade is departmentalized instead of separated or just breaking out for your math or ELA. I love it.”

Teachers also choose practices to manage stress, pressure, and lack of control. Some teachers have specialized to reduce the number of subjects they teach while others have rejected TVAAS as a measure of efficacy. Still others have relocated to untested grades; several of these teachers requested teaching assignments to untested grades within their school or at another school within the district. Other were relocated by their principal to grades, subjects, or positions that are not evaluated with personal TVAAS scores.

“... the teacher would become an expert in that particular subject area because back to the first thing--there's a lot of fear with common core and the standards going so in-depth people were fearful that they could not be an expert on all four subject areas instead of just the one.”

“I mean I don't know, I just, I take, honestly, I take it with a grain of salt. I mean a look at my data, and I think okay, that's interesting that this student did not, you know, but for the most part, I mean it doesn't, is not going to, make me feel bad if I don't do well on the, if I don't have good scores.”

“I took my job very seriously, and I think that's why I wanted to move out of, you know being tested because I felt like I failed a lot of times.”

Situating the Model

The following section situates the theoretical model within grounded theory reliability criteria and within the conceptual framework of this study.

Within grounded theory reliability criteria. This qualitative, situated inquiry does not attempt to make generalizations or make claims of transferability because the phenomena described are intimately tied to the interpretation and context of the participants. However, Glaser and Strauss (2008) offered criteria specifically for evaluating the reliability of a grounded theory. *Reliability*, as used in this qualitative study, is not the same as the criteria for statistical reliability used in quantitative research. Instead, Glaser and Strauss (1968) evaluate reliability in these terms: *fit, work, relevance, and modifiability*.

Fit. Glaser and Strauss suggested that a theoretical model should represent the empirical findings from the data, not a pre-established theoretical perspective. Therefore, the analysis of this study relies on codes created from teachers' responses through constant comparative method rather than preselected codes derived from an existing theory. For member checking, all participants received a copy of their interview transcript and invited to share any concerns or questions with the researcher. None of the participants asked for changes to be made to their answers. Additionally, the four core themes are generated through constant comparative methods and continue to evolve throughout the analysis. Therefore, the teachers' responses provide the context and boundary of the study so that readers can test the "degree of fittingness" (Guba, 1981, p. 86). However, Lomborg and Kirkevold (2003) suggested that from a relativist position,

the *fitness* of a theory cannot be evaluated by its accuracy in representing reality.

Therefore the credibility of the theoretical model of this study depends upon how the data collected represent the meaning that was constructed and reconstructed by the researcher (Lomborg & Kirkevold, 2003).

Work. The theoretical model provides predictions, explanations, and interpretations of the phenomena of this study. The model goes beyond describing the phenomena; it offers a set of relationships among the conditions of participants' lives, their experiences, the strategies they choose and the consequences of those choices. It not only identifies teachers' experiences and practices, but it also suggests constructs to explain the interactive dynamics between them.

Relevance. A model's relevance depends upon its ability to respond to the contexts of the research question and to the central research problem. The strength of this theoretical model is that its implications can be useful to teachers beyond the study's participants. The model can provide insights for other teachers facing value-added evaluation policy, for school and district level administrators who transmit evaluation policy through practice and for decision makers who select or design teacher evaluation policy.

Modifiability. Finally, a theoretical model is also measured by its ability to adapt as new data and situations emerge and still remain intact and relevant. Constructivist grounded theory recognizes social reality as dynamic; therefore, fitness of a model also depends upon its generation from situated, empirical data and its ability to adapt when new data become available or new contexts are introduced.

Within foundational theories and conceptual framework. The following section situates the theoretical model within the foundational theories and conceptual framework of the study.

Within sensemaking theories. Sensemaking theories draw from symbolic interactionism to provide a heuristic framework for understanding how teachers respond to reforms. Blumer (1969) offers three premises:

- people react to things based on the meanings they have ascribed to those things
- people construct meaning of things through their social interactions with others
- people use an interpretive process to mediate the meanings they construct to deal with the things they encounter.

Teachers' reactions to policy are based on the meanings they construct through interactions with others, including teammates, colleagues, and administrators. Coburn (2001) argued that teachers socially construct and reconstruct meaning as they “actively mediate norms, belief systems, and practices that have diffused from the institutional environment” (p. 147). She also posited that teachers make sense of implicit messages of policy through the norms and routines of their departments or workgroups in the organizational environment (2001). Kelchtermans (2009) suggested that sensemaking, an “ongoing interactive process,” (p. 263) can explain how teachers look at their professional situation, give meaning to it, and react to it. Sensemaking is based on a personal system of knowledge and beliefs that Kelchtermans (März et al., 2016) called a

personal interpretative framework that acts as a “cognitive and affective lens” (p. 309) for teachers. Components of the personal interpretative framework include self-image, self-esteem, task perception, job motivation, and future perspective. This framework filters what teachers observe, how they interpret their observations, and how they evaluate reforms.

Findings. Personal interpretative frameworks can be useful for modeling how teachers’ professional histories shape how they make sense of value-added evaluation policy. Teachers’ professional histories are shaped by teachers’ TVAAS scores, their years of evaluation, their years of teaching experience, and their beliefs of self-efficacy. Cultural beliefs and norms, institutional norms and practices, and value-added evaluation policy communicate institutional messages about the control teachers have over student outcomes. Teachers filter these institutional logics through a personal framework of knowledge and beliefs that they develop through personal experiences with value-added evaluation policy, years of teaching, perceptions of self, and what they believe they should be doing for their jobs. Teachers also make sense of these messages through their interactions with colleagues and supervisors as well as their students. When teachers are faced with a new evaluation policy, task perceptions or perceptions of self-efficacy can be challenged by changing institutional beliefs and related practices. This disequilibrium between institutional logics of schools and professional self-understanding of educators can cause teachers to feel as if their beliefs and values are undervalued or under attack. This can result in damaged professional self-esteem and diminished self-efficacy (März, Kelchtermans, & Dumay, 2016).

Kelchtermans' (2009) sensemaking theory can help to explain how institutional norms related to value-added evaluation cause conditions in which teachers feel a lack of agency and experience pressure to improve student outcomes. Teachers may also leave high pressure grades and subjects for other teaching assignments where they experience more success and less pressure to perform on TVAAS. Drawing from sensemaking theories, this study provides empirical data that illuminate how institutional practices that validate value-added evaluation can affect teachers' self-efficacy and self-esteem.

In the next section, situating the theoretical model within theories of inhabited institutionalism and agency explains the process of how teachers have responded to phenomenological experiences of shame, failure, pressure, and stress, and how they constructed new strategies to build beliefs of self-efficacy through their choices in practice.

Within institutional and sociocultural theories. Institutional theory suggests that messages in the school environment, or logics, are adopted through *normative* means as teachers feel pressured to adopt certain practices to maintain legitimacy (Coburn, 2001; Scott, 2005). Complying with the expectations of institutional logics can create conflict for teachers between their beliefs and goals of teaching and learning and institutional requirements of policy (März et al., 2016). The theory of inhabited institutionalism acknowledges the pressure and constraints of messages in the school environment but insists that institutionalism of policy requires agency of individual actors in an institution, not just policy logics (Hallett & Meanwell, 2016; März et al., 2016). Teachers exercise agency and “talk back” through strategies they develop and adapt. Ecological models

used in sociology complement inhabited institutionalism's social constructions of policy implementation. Ecological models focus on people's transactions with their physical and sociocultural surroundings, or "environments," in contrast to behavioral models and theories that do not consider broader community, organizational, and policy influences on behaviors. Sociocultural approaches to understanding teacher responses to reform recognize that the thoughts, beliefs, and actions of individuals are influenced by cultural, historical, and social structures reflected in "mediational tools ... such as policy mandates, curriculum guidelines, and state standards" (Lasky, 2005, p. 900). Social ecological models also challenge the ideology of individual responsibility that belies the complexities of social causation (McLeroy et al., 1988, p. 352).

Findings. Inhabited institutionalism is useful for identifying the embedded ideologies of value-added evaluation policy. Specifically, comparing and contrasting policy logics with the beliefs of teachers helps researchers understand the tension teachers experience when evaluated with value-added methods. Identifying the embedded ideologies of value-added evaluation policy also helps to explain the practices that teachers choose to respond to this tension. Understanding how normative pressures of institutional structures diffuse policy logics within schools is also valuable for understanding the phenomena of stress, pressure, and lack of control over student outcomes. Exploring normative pressures of institutional practices also helps to explain why the teachers in this study chose practices that conform to those logics even if their beliefs disagreed with them.

Tension between logics of value-added evaluation policy and teacher beliefs.

Value-added evaluation policy assumes a direct causal relationship between teacher practices and student growth scores. Sanders and Horn (1994) acknowledge that there are “many factors other than teacher influences that affect student learning,” but they claim that “these influences can be filtered without having to have direct measures of all the concomitant variables“ (pp. 304-305). This assertion places full accountability for student growth scores on individual teachers and incorporates the free market ideology of individual accountability into teacher evaluation.

Findings. Teachers are acutely aware of the vulnerability inherent in teaching and learning relationships, yet with value-added models, teachers are solely responsible for their students’ learning outcomes. As a result, teachers discover their beliefs about their professional vulnerability are largely in conflict with the logics of value-added evaluation policy. Participants disagree with the underlying assumption of TVAAS—that value-added models fully control for all external variables outside of their teaching—because teachers believe many factors outside of their influence still affect their students’ standardized test scores, including their students’ home lives, test-taking abilities, and physiological state during the annual testing period. Teachers also share a variety of reasons why they believe TVAAS is not appropriate for their evaluations. Many argue that TVAAS should not be used to evaluate teachers because it is not a valid measure of teacher efficacy. Some participants express that standardized tests are invalid instruments for measuring what students have learned or what students should be learning. Additionally, teachers argue that student outcomes should not be used to evaluate

teachers because of issues of fairness, the limits of what a standardized test can measure, and the lack of direct causation between teaching and TVAAS scores. Finally, most of the teachers in the study suggest teacher improvement as the best use of evaluations, but they did not find their value-added scores to be helpful for improving their own teaching because the results are returned too late.

Institutional structures that diffuse policy logics. Proponents of the free market paradigm argue that expanding the role of competition and standardized meritocracies in the public sector provides a more efficient means to manage society than policies that seek to lessen income inequality (Tienken, 2013). As a result of the economic and political shifts toward free market ideologies in the public sector, high-stakes standardized testing systems are now used in public schools across the U.S. (Day, 2002). Institutional practices within schools validate standardized value-added scores as measures of teacher efficacy and convey these free market ideologies of competition and standardized meritocracy.

Findings. This study reveals several institutional practices that diffuse value-added policy logics through normative pressures. One example is the annual data meetings at which individual teachers' TVAAS scores are displayed and discussed by peers and administrators within their school. The public nature of these meetings sends a message to communities of teachers that TVAAS scores are not only valid but also the chief source of data for evaluating teachers and schools. The feelings of shame and humiliation experienced and witnessed by the participants at these meetings create powerful social incentives for teachers to meet performance expectations. With TVAAS,

teachers are evaluated using norm-referenced, standardized student data that effectively compares one teacher to another and turns teaching a highly competitive activity. Additionally, the strategic placement of teachers in subjects and grades by principals to improve school TVAAS scores publicly signals the perceived efficacy of teachers based on their contribution to a school's overall performance on a standardized scale. This de facto ranking of teachers introduces the private sector norm of entrepreneurial competition into public school culture.

Participants recount informal conversations, conferences, and formal meetings with principals and district level administrators to focus their instruction and resources on improving the growth scores of certain students and on standardized test scores in general. Through these interactions, principals and other administrators communicate standardized meritocracy as a fair and efficient method of measuring student learning and teacher quality. When outcomes are standardized, the value of both inputs and outputs can be measured and ranked efficiently. School and district administrators' interactions and practices focusing on students' standardized test performances convey a complicity with the commodification of public schools, with teachers as the inputs and students as the outputs.

Within theories of agency. Agency is the capacity to exercise control over one's own thoughts, motivations, and actions. People can effect change in themselves and their situations through their own efforts (Bandura, 1989). The core features of human agency enable people to play a part in their self-development, adaptation, and self-renewal with changing times (Bandura, 2001) Agency is expressed through relationships with others

and social and cultural structures. Through these relationships, environments and people shape each other (Bandura, 2001). To unify the theory of agency as relational and socially situated, social cognitive theory offers a model of triadic reciprocal causation (personal factors, behavior, external environment) to illustrate how action, cognitive, affective, and other personal factors and environmental events interact to direct human action (Bandura, 1989; Wood & Bandura, 1989). Agentic actors reflect on their world, decide when to act upon it, and then reflect on their intervention (Bhaskar, 1989). Agency is the action of individuals based on the meaning they make through interactions with their environment. Consistent with social constructivism and inhabited institutionalism, it provides a framework to explain individuals' behavior based on their sensemaking within social contexts.

Findings. Bandura's (1989) theory of agency claims that belief of self-efficacy is the most powerful mechanism for exercising agency. Many of the teachers in the study reported feelings of failure and shame due to certain institutional practices, including faculty meetings where TVAAS scores were shared with peers. Teachers also reported negative and positive experiences with their principals. Most teachers adopted strategies to influence their students' growth scores. The participants responded to the pressure of cultural and institutional norms through two categories of adaptive strategies:

- striving to meet performance expectations
- managing stress, pressure, and lack of control

These strategies were designed to improve teacher TVAAS scores, manage pressure to improve TVAAS scores, or both. See Table 7 for specific strategies and their intended results.

Table 7. Teachers' strategies and intended results for improving TVAAS scores

Strategies	Intended Intermediate Result	Intended Final Result
Develop students' test taking skills	Improve student performance on standardized tests	Meet performance expectations
Align curriculum and instruction with test content and format		
Increase efficacy through departmentalization		
Individualize skill-based instruction		
Reduce number of subjects taught through departmentalization	Diffuse accountability for students' growth scores	Manage pressure, stress and lack of control
Relocate to grades that are not evaluated on individual growth scores		
Reject TVAAS as a valid measure of efficacy	Maintain belief of self-efficacy	

These strategies are how teachers “talk back” to policy. Teachers believe that TVAAS measures how well students can take standardized tests, not necessarily what teachers have taught or what students have learned. Teachers do not have control over standardized tests, but they do exercise agency to improve students' growth scores by preparing students for tests, aligning curriculum and instruction with the tests, departmentalizing, and individualizing skill-based instruction. Teachers attempt to change their students and themselves by introducing new practices, adapting existing ones, or rejecting policy logic. Teachers have little or no voice in designing evaluation

policy, but they exercise agency through strategies they adopt, adapt, or refuse in response to value-added evaluation policy.

Within vulnerability theories. Palmer (1998) indicated, “As we try to connect ourselves and our subjects with our students, we make ourselves, as well as our subjects, vulnerable to indifference, judgment, ridicule” (p. 17). Kelchtermans’ (2005) theory of teacher identity suggested that vulnerability is a structural condition of teaching. Vulnerability is a structural characteristic of the teaching profession because conditions are out of a teacher’s control, measuring effectiveness is difficult, and the uncertainty of the outcomes of a teacher’s decisions have moral consequences (Uitto et al., 2016). van Veen et al (2005) argued that the teaching relationship is not just technical or intentional; while the instrumental outcomes of teaching are a concern, more is “at stake” (2005, p. 918). Teachers rely upon their values to make decisions that have moral consequences because they affect the lives and needs of children (Kelchtermans, 2005). At the same time, a teacher can never have full command of the outcomes of his/her actions because the pedagogical relationship “radically escapes control and intervention” (Fenstermacher, 1990, p. 998). This relationship of responsibility combined with teachers’ lack of control over the outcomes of teaching creates conditions of ethical vulnerability.

Findings. In contexts of reform, teachers experience vulnerability which mediates their responses to policy (Kelchtermans, 1996; van Veen et al., 2005). Vulnerability as a structural aspect of teaching can help explain the phenomenological responses of teachers to value-added evaluation. Widespread distrust of the validity of TVAAS among the participants reveals that most teachers do not see a direct causation between their efficacy

and student growth scores. Teachers expressed their concerns with the validity of TVAAS scores because of factors out of their control including standardized test conditions, test format, and challenges in their students' lives. Despite claims that TVAAS controls for factors other than a teacher's influence on student performance, nearly all of the teachers in the study experience some lack of control over their students' growth scores. Teachers' feelings of little control over test scores contradicts the policy logics that TVAAS accurately and reliably isolates and measures the efforts of teachers (Fenstermacher, 1990). As discussed earlier, teachers' experiences of vulnerability are in conflict with their perceptions of the validity of value-added evaluation policy, which explains the tensions many participants experience when evaluated with value-added methods. In addition, normative pressures of institutional structures influence teachers' strategies for choosing practices. These pressures explain why teachers choose to strive to improve student growth scores despite their experiences of lack of control and vulnerability. Teachers choose these strategies as a means to reduce their vulnerability and exercise agency over their TVAAS scores.

Within empirical research of value-added evaluation. The following section situates the theoretical model within the empirical research of value-added evaluation including teachers' responses and their consequences and teacher perceptions.

Teachers responses and their consequences. High-stakes value-added evaluation leads to changes in *what* is taught and *how* it is taught in classrooms. In response to evaluation reforms, teachers narrow curricula to match content on standardized tests, devote more instructional time to standardized test preparation and practice, and

substitute constructivist pedagogies with teacher-centered and standardized delivery of content to improve their students' test scores (Au, 2011; Hargreaves, 2005; Hursh, 2007; Nichols & Berliner, 2007). The pressure to produce growth on standardized tests creates a conflict for teachers who must choose between meeting institutional expectations and requirements or the goals and requirements of best practices in teaching and learning (J. W. Meyer & Rowan, 1977; Scott, 2014). Teachers have also become strategic about to which students they devote limited resources. Booher-Jennings (2005) found that in a Texas school district, teachers have engaged in "educational triage," a system of sorting students by their perceived potential to improve their test scores and then allocating more resources for teaching students perceived to be most capable to improve the schools' accountability rating and designating fewer resources for students who are viewed as "hopeless cases" (pp. 232-233).

Findings. Teacher responses in this study echo several behaviors reported in studies within the extant literature on value-added evaluation implementation. Some teachers share with some ambivalence that they kept the test "in mind" when planning and teaching. Many of the participants report that they allocate significant instructional time to prepare students for annual standardized tests. Teachers share a range of practices related to test preparation, including review of test content, test format, teaching testing strategies, conducting practice tests, and building stamina of students to prepare for standardized test conditions. Teachers describe how they use class time to administer practice tests with similar formats and conditions to improve students' performance on the actual tests. Teachers also report planning curriculum and instruction with

standardized test content, format, and language in mind. Several teachers explain how they use their projections of what content the test would include to determine how much instructional time they would spend on a particular set of corresponding standards.

Two practices widely reported in the study that are not found in the literature include individualizing instruction and departmentalizing. Participants who indicate adaptations in teaching most often include individualizing instruction for students to improve their growth. However, several teachers report using individualized instruction to strategically focus teaching resources on “bubble kids,” or students who had the most potential to show growth on standardized tests, similar to the triage approach seen in Texas.

Participants share concerns about producing adequate growth in their students if they are required to teach all subject areas and all express positive emotions when describing their experiences with some form of departmentalizing. Teachers report increased beliefs of self-efficacy when they transition from a self-contained classroom to teaching one or two subjects. They describe themselves as “experts” with a deeper focus on fewer standards and as effective teachers who “fit” well with the subjects they teach.

Teacher perceptions. Overwhelmingly, teachers distrust value-added methods and report low morale as a result of perceived changes in their teaching in order to earn a successful professional rating (Collins, 2014; Feuer, 2012; Firestone, 2014; Goldhaber, DeArmond, & Deburgomaster, 2011; Jiang, Sporte, & Luppescu, 2015; Lee, 2011; Michalek, 2014; Tuytens & Devos, 2009). Teachers are frustrated with the lack of transparency of value-added models, are skeptical of claims that the models can control

for external influences including mobility and poverty, and are doubtful that value-added models can measure the critical aspects of education that cannot be quantified (Jiang et al., 2015; Lee, 2011; National Research Council, 2009).

Findings. Participants' perceptions of TVAAS were consistent with the literature on teachers' responses to other value-added evaluation policies. All 19 of the participants share concerns with the validity or application of TVAAS in their professional evaluations. They describe TVAAS as a "waste of time," "a corrupt system," "too narrow," and "not accurate or fair." Several teachers feel that their students' growth is not accurately reflected in their test scores due to factors outside of their control, including their students' low socioeconomic status and challenges they face at home, the differences in rigor between tests across grade levels, and the inability of TCAP tests to capture qualitative aspects of their students' learning. Teachers also share their concerns with the stress surrounding high stakes testing in their classrooms and the emotional duress it causes their students. Several participants comment that TCAP tests are developmentally inappropriate for elementary age students and that the validity of the tests is dubious because of the negative effects of students' stress on their TCAP test performances. Finally, several participants echo the literature's evidence of confusion among teachers about value-added models. Teachers in the study admit to not understanding how TVAAS works, describing the value-added "formula" as "secretive" and a "great myth."

Implications and Recommendations

Limitations of the Study

The researcher conducted an extensive review of extant literature on teachers' responses to value-added evaluation and found a limited number of qualitative studies on teachers' processes for adapting practice. The study is designed to provide an approach to the study of teachers' responses to policy that advocates sensemaking, a process based on a personal system of knowledge and beliefs, to explain how teachers make choices in their practices. The study required 19 interviews to reach theoretical saturation. Due to the small sample size and qualitative approach of this study, the findings cannot be generalized or scaled-up. This research presents voices of teachers about their lived experiences with policy and how they make sense of that policy through the forms of agency available to them. Although it was beyond the scope of this study, future work could include sharing the findings and analysis with the participants along with an invitation to individually discuss the resonance and relevance of the theoretical model. Focus group interviews with participants could also provide opportunities for participants to evaluate how well the data and analysis represent the meanings they intended. Giving voice to participants regarding the model's credibility would inform revisions and improving the quality of the research.

Delimitations

Due to the methodology, focus, and setting of this study, the findings cannot be generalized. The theoretical model is ecologically situated and cannot be transferred beyond the context where it was generated. The study's focus was limited to a single

policy for a subset of elementary education teachers, and the setting was limited to a single school district. The researcher conducted the study in the district where she has taught for the last seven years. Therefore, while the participants were chosen to reach theoretical saturation, they were selected from a population of teachers the researcher knew through professional networks. A different theory may have been reached with a different set of teachers.

Implications for Practice

Conclusion 1. *Through their choices in institutional practices, school and school district administrators can exercise agency in how policy is implemented.* Cultural beliefs about the role of teachers, value-added evaluation policy, and the structural conditions of vulnerability in teaching can cause teachers to experience tensions between performance expectations and their feelings of lacking control over their students' growth scores. Institutional practices diffuse cultural messages about education that can exacerbate these tensions or mitigate them. Leaders within schools and school districts are often unaware of the hegemonic, neoliberal, and market-based ideologies in education policy that institutional practices transmit. While the causal conditions of these phenomenological tensions of teachers are beyond the immediate control of schools and school districts, school norms and practices *can* influence teachers' responses to policy. Teacher agency is mediated by their social contexts; institutional practices and norms shape teachers' choices of practice including displays of TVAAS scores at data meetings, strategic placement of teachers, leveling students by ability, and departmentalizing. Therefore, schools and districts would benefit from considering the potential impact of institutional

practices on teacher agency when choosing institutional practices to implement new policies.

Conclusion 2. *As teachers implement value-added evaluation policy through their responses to the policy, the consequences of those responses can change teaching practices in unanticipated ways.* In response to the phenomena, elementary education teachers are choosing several strategies to improve student test scores. Teachers are using instructional time with students to develop test-taking skills and aligning curricula and instruction with the test content and format. These practices lead to several consequences in schools. One outcome of value-added policy is that it creates incentives for teachers to pass over students that are deemed to have the least potential to achieve proficiency. Teachers respond to value-added evaluation policy by funneling resources to “bubble kids,” the students that teachers believe can show growth and can “move” from one standardized achievement test ranking to the next. Teachers re-direct their efforts to “bubble kids” at the expense of others because these students have the greatest potential to impact teachers’ TVAAS scores. Therefore, while the stated purpose of using student test scores in teacher evaluations is to improve teacher effectiveness, teachers are responding with strategic practices that can boost scores but don’t necessarily improve teaching efficacy.

Teachers are also devoting planning and instructional time to individualizing skill-based instruction. Teachers report being more aware of students’ progress towards mastery of standards when evaluated with TVAAS. In response, teachers are designing their instruction to address deficiencies of mastery. Schools, teaching teams, and teachers

are also utilizing teaching assistants to address students' deficiencies through small group interventions. Teachers are also planning and collaborating curricula with their colleagues both formally and informally to encourage students to perform better on standardized tests.

Elementary education teachers are specializing in subject areas to build beliefs of self-efficacy in two ways. Teachers perceive that teaching fewer subject areas means fewer standards to address and will require less planning effort and time. Given the limited time teachers have to prepare lessons, having fewer standards means teachers have more time per standard to prepare quality lessons to teach the standards for which they are responsible. Teachers express increased self-efficacy as a result of more preparation time per standard. Teachers also express an increase in self-efficacy when they teach in the subject areas of their strength. Teachers perceive that their personal aptitudes, comfort levels, and enthusiasm for preferred subject areas increase their efficacy to reach students, to help them learn more effectively, and to increase their TVAAS scores. The alignment of teachers' belief of self-efficacy with their expertise of standards in tested content areas reflects Freire's banking model of education where teachers make deposits of knowledge, rendering students as passive and uncritical receivers of information and skills.

Elementary education teachers are creating new ways to exercise agency over student growth scores and to manage stress and pressure of meeting performance expectations. However, the effect of specialization of teachers on the lived experiences of elementary school students is largely unknown. Additionally, the impact of shifts in how

students' instructional time is spent, in the content of curriculum and in pedagogy has yet to be fully examined. The current focus in classrooms is on mastery of standards, but this is perhaps at the expense of educating "the whole child, developing learners who are knowledgeable, emotionally and physically healthy, civically active, artistically engaged, prepared for economic self-sufficiency, and ready for the world beyond formal schooling" (ASCD, 2011).

Conclusion 3. *Teachers "talk back" to institutional structures by reconstructing norms and practices of schools and districts through their choices in practice.* Teachers make sense of policy as individuals by filtering their experiences through their unique professional history. However, contextual conditions limit teachers' agency in choosing their responses to policy. Consequently, many teachers in the study choose similar strategies for exercising agency over student outcomes and reducing pressure to improve student outcomes. Collectively, teachers' individual responses to value-added evaluation can alter the norms and practices across a school or even a district, creating a positive feedback loop that amplifies these unanticipated shifts.

Schools and districts should pay particular attention to how teachers' responses to value-added evaluation policy are shaping norms and practices, especially collaboration among teachers and using TVAAS scores for hiring and tenure decisions. Schools and school districts should consider the impact of value-added evaluation policy on collaboration within and across schools and its impact on school and district improvement plans. Schools and school districts should also consider potential changes in hiring decisions by principals. As TVAAS scores and expertise in subject areas become

more valued within school norms, they could play a larger role in determining the quality of teacher candidates. Changes in criteria for teacher effectiveness can significantly change a faculty of a school over time as new teachers with expertise in one or two subject areas are hired over teachers with more less-specialized experience. Because the consequences of teachers' responses to policy can become new norms for schools across the district, notice should be paid to patterns in adaptations of teachers' practice.

Implications for Policy

Conclusion 4. *Value-added evaluation policy creates a tension between cultural ideologies and teachers' lived experiences of vulnerability.* Value-added evaluation policy reifies several market-based beliefs about education. The TVAAS is built upon data that indicates a direct causal relationship between teacher practices and student learning. TVAAS scores as a component of teacher evaluation imply that consequences (or incentives) are required for teachers to engage in effective practices. Additionally, TVAAS legitimizes the standardization of meritocracy because it is calculated from standardized test scores that are accepted as fair and efficient measures of student learning and teacher quality.

Value-added evaluation policy asserts that teachers have full control over their students' TVAAS scores because the value-added algorithm eliminates all extraneous factors on student learning, including socio-economic conditions, race, class, etc. This leaves only the teacher's performance as a determinant of student outcomes on standardized test. However, the empirical data collected in this study overwhelmingly suggest that most teachers experience limited capacity to exercise agency over student

outcomes. Teachers believe that they influence their students' learning, but they also know that pedagogically they can never fully control the outcomes. Teachers also perceive the validity of standardized measures of student learning to be flawed in practice because testing conditions are never fully controlled. As a result, value-added evaluation policy creates a paradox of agency for elementary education teachers in this study.

Denzin (2001) suggested that those who design and implement applied policy "must grasp, interpret, and understand the perspectives and experiences of those persons they intend to serve if their programs are to be both solid and effective" (p. 120). Policy makers must consider the vulnerability inherent in the experiences of teachers when designing evaluation policies. Evaluation policy must address how teachers experience a lack of control or ability to exercise agency over student outcomes. Additionally, policy makers would benefit from exploring the shortcomings of standardized measures of student outcomes in practice. Teachers' experiences with state standardized tests reveal flaws in validity and problematizing these flaws can inform the design of improved teacher evaluation and accountability systems.

Suggestions for Future Research

Given that the field of inhabited institutionalism is relatively new and empirical data on the theory are limited, additional studies to explore institutional processes in school contexts may provide meaningful conclusions to increase understanding of implementation of educational policy. In addition, a plethora of academic research and literature on teachers' responses to value-added evaluation exists, but empirical data on

how teachers choose their responses are lacking. The perception of a problem defines the response to it. A process-oriented approach can help identify potential unintended consequences of policy for both schools and districts. Finally, given that teachers' practices are adapting to logics of teaching and learning, the impact of value-added evaluation is far reaching and warrants further study. Below is a list of questions that may guide further research into value-added evaluation, inhabited institutionalism, and teacher agency.

Value-Added Evaluation

- With the specialization of teachers and teacher teams departmentalizing in elementary schools, what is the impact on the lives of children in schools, and how are non-academic areas of their education affected?
- Is departmentalization a practice adopted beyond the setting of this study? Is it seen across Tennessee and other states with value-added evaluation policies?
- Are pre-service teacher education programs responding to TVAAS with changes in curricula or certification requirements?
- How do pre-service teachers make sense of value-added evaluation policy, and how does it shape their future plans for teaching?
- How has value-added evaluation shaped teachers' perceptions of collaboration with colleagues and other schools?

Inhabited Institutionalism

- How have principals and school district administrators adapted their practices in response to value-added evaluation?
- What other “pressures” do teachers experience in response to evaluation policy?
- What are the institutional logics of the other components of teachers’ evaluation, including observation scores from evaluators, professionalism rubric, etc.?
- Community School District released a new motto in the fall of 2017: “Care, Love, Service,” along with an emphasis on relationships with students. Are institutional logics changing in the district? What institutional practices support or undermine the logics of this motto?

Teacher Agency

- How do teachers experience agency in their evaluations?
- How do teachers experience vulnerability in their daily lives?
- How can school practices address the paradox of agency and value-added evaluation?

Chapter Summary

This study responds to the empirical research questions by describing and interpreting the lived experiences and the choices of practice of teachers evaluated by the Tennessee Value-Added Assessment System. The research findings are used to construct

a theoretical model to explain how teachers choose their practices in the context of their experiences. The discussion focuses on situating the theoretical model within reliability criteria and the conceptual framework. The study is limited by the sample size and qualitative approach. Thus, the findings cannot be generalized or scaled-up. Delimitations of the study include its grounded theory methodology, its focus of a single policy, and its setting within a single school district. The participants were selected from a population of teachers the researcher knew through professional networks, so it is possible that the sample is skewed toward a particular theory.

The implications for practice and policy offer four conclusions. First, administrators can and do exercise agency in how policy is implemented through their choices in institutional practices. Second, teachers' responses to value-added evaluation can change teaching practices in unanticipated and unintended ways. Third, teachers "talk back" to institutional structures by reconstructing norms and practices of schools and districts through their choices in practice. Lastly, value-added evaluation policy creates a tension between cultural ideologies and teachers' lived experiences of vulnerability and agency. In the discussion of implications for further research, questions to explore regarding value-added evaluation, inhabited institutionalism, and teacher agency are suggested.

Researcher's Reflection on Learning Journey

Although I had no intention of teaching, my senior thesis for my undergraduate degree in the 1990s was on the effects of decentralization of schools in the Los Angeles Unified School District on school accountability. As a Politics, Philosophy and

Economics major, education policy stood perfectly at the intersection of power and theory, with money always standing somewhere close by. I blame my initial rejection of teaching as a profession on my mother, an incredibly talented educator who changed dozens of lives as a literacy coach and curriculum administrator for a large school district in Illinois. She and my father were both civil servants and because of their high opinions of me and their frustrating experiences with government bureaucracies, they argued that I would waste my talents if I went into teaching. Following their encouragement, I went into professional consulting after earning my BA, imagining it to be an intellectually stimulating and financially rewarding career. It was both challenging and lucrative, but I could not find a greater purpose in the work. The long hours and stressful deadlines were in the service of for-profit corporations, and I found the work meaningless. I struggled to discover what contributions my labor could make to improve the lives of others as my mother's had.

Within six years, I returned to school for a Masters of Arts in teaching elementary education through an alternative certification program. This meant I could teach, but with minimal formal education in education. Fortunately, I began my teaching career at a laboratory school where the exchange of practice and theory was part of my everyday experiences. My intellectual curiosity and my desire to learn more about practical and theoretical teaching and learning were met constantly with opportunities to participate in research, to design and implement curricula, and to collaborate with other progressive educators. I lived in an ivory tower with ambitions of improving the lives of children and teachers by one day joining the ranks of my mother in public education.

That opportunity came after three years of teaching full-time and four more years of staying home to raise my two sons. I was hired to teach at a high-poverty Title I school in Tennessee. I had high expectations, and I approached the opportunity as more than a position; I was entering a vocation. However, my first year of teaching at a public school after a four-year hiatus from a private, progressive laboratory school, was at best disappointing and at worst disillusioning. A new state-mandated evaluation system had been implemented, and the responses of teachers and administrators created an environment of fear, powerlessness, and distrust.

The emotionally meaningful experiences in my first year of teaching at that public school planted the seeds of this study. New to the school and system, I was an outsider trying to make sense of the responses and dynamics between teachers and the principal. As a newly hired teacher without tenure, I was an insider as well, trying to navigate the evaluation system for my own professional success and security. Over time as I became more integrated as a member of the school system, my shock and fear transformed into anger towards the logics of accountability, fidelity to scripted instruction, and constraints on teacher agency. Meanwhile, my desire to continue my formal education of education led me to enter a doctoral program.

If the emotions of my first year of public school teaching were the seeds of this study, it was my doctoral program that provided the intellectually safe environment for those seeds to germinate. With the support of my peers and professors, I attempted to make sense of the fear, shame, and lack of control teachers experience by asking questions of why teachers have these responses to value-added evaluation. When I first

encountered Freire's emancipatory pedagogies early in my doctoral studies, I heard his call to work critically inside of schools and to collaborate strategically with social movements outside of the education system (P. Mayo, 1999). Finding a vocation in public school teaching and seizing the opportunities to pursue critical and social theory as a graduate student also provided me with a venue to unpack my frustration and anger and begin to question the larger, hidden messages around me that had created my earlier sense of disillusionment. Knowledge is the greatest source of agency for me, and my doctoral studies were a factor in my decision to continue teaching despite my disappointment and to continue the search for opportunities to impact the lives of teachers and students for a greater purpose.

The journey to earning the doctorate has been six years of balancing family, teaching, and graduate studies. Moving forward has been a struggle periodically, and the lack of time has been frustrating, but I am grateful to be able to look back on the process and say that the time was well spent. From the start, I decided that the dissertation would not be a task to complete but rather a process of personal transformation and growth. Given my life's circumstances, I knew the unlikelihood that my dissertation would transition me from practice into full time research. Instead, my goal was to learn and integrate a variety of new analytic and synthetic tools into my role as a teacher and to communicate ideas that can catalyze change within an educational organization. Because transformation is a process, not a task, I had to give up control of the timeline that the degree would require, and the professional growth I've experienced along the way was worth the wait.

The participants in my study have been an influential force in this transformational learning experience. I have been humbled by their passion for their students' well-being and their deep understanding of teaching and learning. Each of these teachers led me through a part of this learning journey, and I am no longer in the same place I started because of their courage to share their stories. Because of the collective knowledge of the teachers in my study, I have a better understanding of how people come to "know" what they know. They have given me a broader perception of agency and taught me to recognize it more readily in the activities of teachers. They have also taught me how to better empathize with fellow teachers, especially when we disagree philosophically. Their willingness to be vulnerable about their experiences encouraged me to do the same, and as a result, I have come to know myself better including my insecurities and biases. I hope to impart the gifts of these transformations with others sharing this profession that often feels more like a calling. And to those outside of the classroom, my goal is to communicate my emic perspective of teaching and to amplify the voices of teachers faced with policies that belie the complexity and vulnerability inherent in teaching and learning.

REFERENCES

- Acker, S. (1995). Carry on caring: The work of women teachers. *British Journal of Sociology of Education, 16*(1), 21–36.
- AERA Council. (2015). AERA statement on use of value-added models (VAM) for the evaluation of educators and educator preparation programs. *Educational Researcher, 44*(8), 448–452. <https://doi.org/10.3102/0013189X15618385>
- Aldrich, M. W. (2017, March 20). William Sanders, pioneer of controversial value-added model for judging teachers, dies [Blog post]. Retrieved May 19, 2018, from <https://www.chalkbeat.org/posts/tn/2017/03/20/william-sanders-pioneer-of-controversial-value-added-model-for-judging-teachers-dies/>
- Allman, P., & Wallis, J. (1992). Radical adult education: Back to basics. In N. Miller & L. West (Eds.), *Changing cultures and adult learning* (pp. 11–13). Canterbury, England: University of Kent.
- Allman, P., & Wallis, J. (1995). Challenging the postmodern condition: Radical adult education for critical intelligence. In M. Mayo & J. Thompson (Eds.), *Adult Learning, Critical Intelligence and Social Change*. Leicester, England: National Inst. of Adult Continuing Education. Retrieved from <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED396138&site=eds-live&scope=site>
- Amrein-Beardsley, A. (2008). Methodological concerns about the Education Value-Added Assessment System. *Educational Researcher, 37*(2), 65–75. <https://doi.org/10.3102/0013189X08316420>

- Amrein-Beardsley, A. (2018, April 26). An important but false claim about the EVAAS in Ohio [Blog post]. Retrieved from <http://vamboozled.com/an-important-but-false-claim-about-the-evaas-in-ohio/>
- Anderies, J. M., & Janssen, M. A. (2013). Robustness of social-ecological systems: Implications for public policy. *Policy Studies Journal*, *41*(3), 513–536.
- Anfara, V. A. J., Brown, K. M., & Mangione, T. L. (2002). Qualitative analysis on stage: Making the research process more public. *Educational Researcher*, *31*(7), 28–38. <https://doi.org/10.3102/0013189X031007028>
- Apple, M. W., & Jungck, S. (1990). “You don’t have to be a teacher to teach this unit”: Teaching, technology, and gender in the classroom. *American Educational Research Journal*, *27*(Summer), 227–251.
- Aronowitz, S. (2003). *How class works: Power and social movement*. New Haven, CT: Yale University Press.
- ASCD. (2011, August 15). ASCD’S Whole Child Initiative launches map of outstanding whole child examples. Retrieved from http://link.galegroup.com.ezproxy.mtsu.edu/apps/doc/A264304124/AONE?u=tel_middleten&sid=AONE&xid=025de724
- Au, W. (2011). Teaching under the new Taylorism: High-stakes testing and the standardization of the 21st century curriculum. *Journal of Curriculum Studies*, *43*(1), 25–45. <https://doi.org/10.1080/00220272.2010.521261>
- Ayers, W. (2001). *To teach: The journey of a teacher* (2nd ed). New York, NY: Teachers College Press.

- Baker, B. D., Oluwole, J. O., & Green, P. C., III. (2013). The legal consequences of mandating high stakes decisions based on low quality information: Teacher evaluation in the Race-to-the-Top era. *Education Policy Analysis Archives*, 21(4/5), 1–68.
- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18(2), 215–228.
<https://doi.org/10.1080/0268093022000043065>
- Bandura, A. (1989). Human agency in social cognitive theory. *American Psychologist*, 44(9), 1175–1184.
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52(1), 1.
- Bandura, A., & Jourden, F. J. (1991). Self-regulatory mechanisms governing the impact of social comparison on complex decision making. *Journal of Personality and Social Psychology*, 60(6), 941–951. <https://doi.org/10.1037/0022-3514.60.6.941>
- Bartz, D. E. (2016). Revisiting James Coleman's epic study entitled Equality of Educational Opportunity. *National Forum of Educational Administration & Supervision Journal*, 34(4), 1–10.
- Basu, A., & Dutta, M. J. (2008). Participatory change in a campaign led by sex workers: Connecting resistance to action-oriented agency. *Qualitative Health Research*, 18(1), 106–119.

- Bemelmans-Videc, M.-L. (2011). Policy instrument choice and evaluation. In M.-L. Bemelmans-Videc, R. C. Rist, & E. O. Vedung (Eds.), *Carrots, sticks, and sermons: Policy instruments and their evaluation*. Transaction Publishers.
- Berlak, H. (2000). Cultural politics, the science of assessment and democratic renewal of public education. In A. Filer (Ed.), *Assessment: Social Practice and Social Product* (pp. 189–207). London: Routledge. Retrieved from <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=82985&site=eds-live&scope=site>
- Berliner, D., & Biddle, B. (1996). In defense of schools: The authors of Manufactured Crisis challenge education critics and defend the state of U.S. public school system by debunking several education “myths.” *Vocational Education Journal*, 71(3), 36–38.
- Bhaskar, R. (1989). *The possibility of naturalism: A philosophical critique of the contemporary human sciences*. New York, NY: Harvester Wheatsheaf. Retrieved from <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=brb&AN=68574213&site=eds-live&scope=site>
- Bill and Melinda Gates Foundation. (2010). *Learning about teaching: Initial findings from the measures of effective teaching project*. Seattle, WA: Bill & Melinda Gates Foundation.
- Blumer, H. (1969). *Symbolic interactionism: Perspective and method*. Englewood Cliffs, N.J.: Prentice-Hall. Retrieved from

<https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat00263a&AN=mts.b1192101&site=eds-live&scope=site>

- Booher-Jennings, J. (2005). Below the bubble: “Educational triage” and the Texas accountability system. *American Educational Research Journal*, 42(2), 231–268.
- Bourdieu, P. (2000). *Pascalian meditations*. Stanford, CA: Stanford University Press.
- Braun, H. I. (2005). *Using student progress to evaluate teachers: A primer on value-added models* (p. 20). Princeton, NJ: Educational Testing Service. Retrieved from <https://eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED529977>
- Braun, H. I. (2015). The value in value added depends on the ecology. *Educational Researcher*, 44(2), 127–131. <https://doi.org/10.3102/0013189X15576341>
- Bray, L. E., & Russell, J. L. (2016). Going off script: Structure and agency in individualized education program meetings. *American Journal of Education*, 122(3), 367–398.
- Briggs, D., & Domingue, B. (2011). *Due diligence and the evaluation of teachers: A review of the value-added analysis underlying the effectiveness rankings of Los Angeles Unified School District Teachers by the “Los Angeles Times”* (p. 37). Boulder, CO: University of Colorado at Boulder, National Education Policy Center. Retrieved from <https://eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED516008>
- Brockmeier, J. (2009). Reaching for meaning human agency and the narrative imagination. *Theory & Psychology*, 19(2), 213–233.

- Brown, K., & Westaway, E. (2011). Agency, capacity, and resilience to environmental change: Lessons from human development, well-being, and disasters. *Annual Review of Environment and Resources*, 36(Annual), 321–342.
- Bryant, A., & Charmaz, K. (Eds.). (2007). *The Sage handbook of grounded theory*. Thousand Oaks, CA: Sage Publications.
- Buechner, F. (1993). *Wishful thinking: A seeker's ABC* (Rev. and expanded [ed.]). San Francisco, CA: Harper San Francisco.
- Burger, J. M. (1987). *Teacher evaluation policy implementation* (p. 210). Calgary, Alberta: Alberta Dept. of Advanced Education. Retrieved from <https://eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED295303>
- Campbell, D. T. (1976). *Assessing the impact of planned social change. Occasional Paper Series, #8* (p. 74). Washington, DC: National Science Foundation.
- Carley, K. (1993). Coding choices for textual analysis: A comparison of content-analysis and map analysis. *Sociological Methodology*, 23, 75–126.
- Charmaz, K. (2005). Applications for advancing social justice studies. In Norman K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed.). Thousand Oaks, CA: Sage Publications.
- Charmaz, K. (2014). *Constructing grounded theory* (2 ed.). Thousand Oaks, CA: Sage Publications.
- Chase, S. E. (2005). Narrative inquiry: Multiple lenses, approaches, voices. In Norman K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed.). Thousand Oaks: Sage Publications.

- Choi, K., Goldschmidt, P., & Yamashiro, K. (2005). Exploring models of school performance: From theory to practice. *Yearbook (National Society for the Study of Education)*, 104(2), 119–146. <https://doi.org/10.1111/j.1744-7984.2005.00028.x>
- Chubb, J. E., & Moe, T. M. (1990). *Politics, markets, and America's schools*. Washington, DC: Brookings Institution. Retrieved from <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat00263a&AN=mts.b1327057&site=eds-live&scope=site>
- Clarke, A. E. (2003). Situational analyses: Grounded theory mapping after the postmodern turn. *Symbolic Interaction*, 26(4), 553–576. <https://doi.org/10.1525/si.2003.26.4.553>
- Coburn, C. E. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis*, 23(2), 145–170.
- Coburn, C. E. (2016). What's policy got to do with It? How the structure-agency debate can illuminate policy implementation. *American Journal of Education*, 122(3), 465–475.
- Cochran-Smith, M., Carney, M. C., Keefe, E. S., Burton, S., Wen-Chia Chang, Fernandez, M. B., ... Baker, M. (2018). *Reclaiming accountability in teacher education*. New York, NY: Teachers College Press.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. (1966). *Equality of educational opportunity* (Government No. OE-38001) (p. 746). Washington, DC: National Center for Educational Statistics.

- Collins, C. (2014). Houston, we have a problem: Teachers find no value in the SAS Education Value-Added Assessment System (EVAAS®). *Education Policy Analysis Archives*, 22(98), 1–39.
- Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Los Angeles, CA: Sage Publications.
- Corcoran, S. P., Jennings, J. L., & Beveridge, A. A. (2011). *Teacher effectiveness on high- and low-stakes tests*. Evanston, IL: Society for Research on Educational Effectiveness. Retrieved from <https://eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED518223>
- Corcoran, S. P. (2015). Potential pitfalls in the use of teacher value-added data. In J. A. Grissom & P. Youngs (Eds.), *Improving teacher evaluation systems: Making the most of multiple measures* (pp. 51–62). New York, NY: Teachers College Press.
- Creswell, J. W. (2012). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Los Angeles, CA: Sage Publications.
- Darling-Hammond, L. (1991). The implications of testing policy for quality and equality. *The Phi Delta Kappan*, 73(3), 220–225.
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E. H., & Rothstein, J. (2011). *Getting teacher evaluation right: A background paper for policy makers*. Washington, DC: National Academy of Education.
- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating teacher evaluation. *The Phi Delta Kappan*, 93(6), 8–15.

- Day, C. (2002). School reform and transitions in teacher professionalism and identity. *International Journal of Educational Research*, 37(8), 677–692.
- Denzin, Norman K. (2001). *Interpretive interactionism* (Vol. 16). Thousand Oaks, CA: Sage Publications. Retrieved from <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat00263a&AN=mts.b3267621&site=eds-live&scope=site>
- Denzin, Norman K., & Lincoln, Y. S. (Eds.). (2005). *The Sage handbook of qualitative research* (3rd ed). Thousand Oaks, CA: Sage Publications.
- Denzin, Norman K., & Lincoln, Y. S. (2008). *Strategies of qualitative inquiry* (3rd ed). Thousand Oaks, CA: Sage Publications.
- Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. New York, NY: The Macmillan Company.
- Di Placito, M. L. (2011). Considering the methodological complexities of the interview approach when exploring the work of mothering and disability. *Critical Disability Discourse/Discours Critiques Dans Le Champ Du Handicap*, 3. Retrieved from <https://cdd.journals.yorku.ca/index.php/cdd/index>
- Driscoll, M. P. (2005). *Psychology of learning for instruction* (3rd ed.). Boston, MA: Pearson Allyn and Bacon.
- DuFour, R., DuFour, R. B., & Eaker, R. E. (2008). *Revisiting professional learning communities at work: New insights for improving schools*. Bloomington, IN: Solution Tree.

- Eaker, R., & Keating, J. (2011). *Every School, Every Team, Every Classroom: District Leadership for Growing Professional Learning Communities*. Bloomington, IN: Solution Tree.
- Editorial Projects in Education Research Center. (2004). Issues A-Z: Accountability. In *Education Week*.
- Edman, G. W. (2005). *Mediating language, literacy and bilingualism: Community liaisons as cultural and linguistic brokers* (Doctoral dissertation). Retrieved from <http://search.proquest.com.ezproxy.mtsu.edu/pqdtft/docview/305025936/abstract/107078FF682142E2PQ/1?accountid=4886>
- Emirbayer, M., & Mische, A. (1998). What is agency? *American Journal of Sociology*, *103*(4), 962–1023.
- Ewing, J. (2011, May). Mathematical intimidation: Driven by the data. Notices of the American Mathematical Society. Retrieved from https://www.washingtonpost.com/blogs/answer-sheet/post/leading-mathematician-debunks-value-added/2011/05/08/AFb999UG_blog.html?utm_term=.5f14797f81e7
- Fenstermacher, G. (1990). Some moral considerations on teaching as a profession. In J. I. Goodlad, R. Soder, & K. A. Sirotnik (Eds.), *The moral dimensions of teaching*. San Francisco, CA: Jossey-Bass Publishers.
- Feuer, M. J. (2012, September 26). Measuring accountability when trust is conditional. *Education Week*, *32*(5), 32–27.

- Firestone, W. A. (2014). Teacher evaluation policy and conflicting theories of motivation. *Educational Researcher*, 43(2), 100–107.
- Fitz, J. (1994). Implementation research and education policy: Practice and prospects. *British Journal of Educational Studies*, 42(1), 53–69.
<https://doi.org/10.1080/00071005.1994.9973983>
- Flores, M. A., & Day, C. (2006). Contexts which shape and reshape new teachers' identities: A multi-perspective study. *Teaching and Teacher Education*, 22(2), 219–232. <https://doi.org/10.1016/j.tate.2005.09.002>
- Foley, G. (1994). Adult education and capitalist reorganisation. *Studies in the Education of Adults*, 26(2), 121–143. <https://doi.org/10.1080/02660830.1994.11730602>
- Freire, P. (1970a). *Cultural action for freedom*. Cambridge, MA: Harvard Educational Review and Center for the Study of Change and Development.
- Freire, P. (1970b). *Pedagogy of the oppressed*. New York, NY: Herder and Herder.
- Fried, R. L. (1995). *The passionate teacher: A practical guide*. Boston, MA: Beacon Press.
- Friedman, M., & Friedman, R. D. (1980). *Free to choose: A personal statement* (1st ed). New York: Harcourt Brace Jovanovich.
- Fuhrman, S. H. (1999). *The new accountability* (Policy Brief). Philadelphia, PA.: Consortium for Policy Research in Education. Retrieved from <http://www.upenn.edu/gse/cpre/>

- Gabriel, R., & Lester, J. N. (2013). Sentinels guarding the grail: Value-added measurement and the quest for education reform. *Education Policy Analysis Archives*, 21(9), 1–30.
- Geijsel, F., Slegers, P., van den Berg, R., & Kelchtermans, G. (2001). Conditions rostering the implementation of large-scale innovation programs in schools: Teachers' perspectives. *Education Administration Quarterly*, 37(1), 130–166.
- Genat, B. (2009). Building emergent situated knowledges in participatory action research. *Action Research*, 7(1), 101–115.
- Gibson, R. J., & Ross, E. W. (2007). Introduction. In E. W. Ross & R. J. Gibson (Eds.), *Neoliberalism and education reform*. Cresskill, NJ: Hampton Press.
- Giddens, A. (1979). *Central problems in social theory*. London: Macmillan Education UK. https://doi.org/10.1007/978-1-349-16161-4_3
- Giorgi, A. (2006). Difficulties encountered in the application of the phenomenological method in the social sciences. *Análise Psicológica*, 24(3), 353–361.
- Giroux, H. A., & Giroux, S. S. (2006). Challenging neoliberalism's new world order: The promise of critical pedagogy. *Cultural Studies/Critical Methodologies*, 6(1), 21–32. <https://doi.org/10.1177/1532708605282810>
- Glaser, B. G. (1965). The constant comparative method of qualitative analysis. *Social Problems*, 12(4), 436–445.
- Glaser, B. G. (1978). *Theoretical sensitivity: Advances in the methodology of grounded theory*. Mill Valley, CA: Sociology Press.

- Glaser, B. G., & Strauss, A. I. (1968). *The discovery of grounded theory: Strategies for qualitative research*. London, England: Weidenfeld and Nicolson.
- Golden, S. D., & Earp, J. A. L. (2012). Social ecological approaches to individuals and their contexts: Twenty years of health education and behavior health promotion interventions. *Health Education & Behavior, 39*(3), 364–372.
<https://doi.org/10.1177/1090198111418634>
- Golden, S. D., McLeroy, K. R., Green, L. W., Earp, J. A. L., & Lieberman, L. D. (2015). Upending the social ecological model to guide health promotion efforts toward policy and environmental change. *Health Education & Behavior, 42*(1), 8S-14S.
<https://doi.org/10.1177/1090198115575098>
- Goldhaber, D., DeArmond, M., & Deburgomaster, S. (2011). Teacher attitudes about compensation reform: Implications for reform implementation. *Industrial & Labor Relations Review, 64*(3), 441–463.
- Grissom, J. A., Loeb, S., & Doss, C. (2015). The multiple dimensions of teacher quality. In J. A. Grissom & P. Youngs (Eds.), *Improving teacher evaluation systems: Making the most of multiple measures* (pp. 37–50). New York, NY: Teachers College Press.
- Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communication & Technology, 29*, 75–91.
- Hagstrom, M. (2015, April 22). How technology can help elevate the human condition [Blog post]. Retrieved from <https://www.weforum.org/agenda/2015/04/how-technology-can-help-elevate-the-human-condition>

- Hallett, T. (2010). The myth incarnate: Recoupling processes, turmoil, and inhabited institutions in an urban elementary school. *American Sociological Review*, 75(1), 52–74.
- Hallett, T., & Meanwell, E. (2016). Accountability as an inhabited institution: Contested meanings and the symbolic politics of reform. *Symbolic Interaction*, (3), 374. <https://doi.org/10.1002/symb.241>
- Hallinger, P., Heck, R. H., & Murphy, J. (2014). Teacher evaluation and school improvement: An analysis of the evidence. *Educational Assessment, Evaluation and Accountability*, 26(1), 5–28. <https://doi.org/10.1007/s11092-013-9179-5>
- Hanushek, E. A. (2014). Boosting teacher effectiveness. In C. E. Finn & R. Sousa (Eds.), *What lies ahead for America's children and their schools* (pp. 23–35). Stanford, CA: Hoover Institution Press.
- Hargreaves, A. (1998). The emotional practice of teaching. *Teaching and Teacher Education*, 14(8), 835–854. [https://doi.org/10.1016/S0742-051X\(98\)00025-0](https://doi.org/10.1016/S0742-051X(98)00025-0)
- Hargreaves, A. (2000). Mixed emotions: Teachers' perceptions of their interactions with students. *Teaching and Teacher Education*, 16(8), 811–826.
- Hargreaves, A. (2005). Editorial Statement. *The Educational Forum*, 69(2), 101–110. <https://doi.org/10.1080/00131720508984674>
- Harris, D. N., Ingle, W. K., & Rutledge, S. A. (2014). How teacher evaluation methods matter for accountability: A comparative analysis of teacher effectiveness ratings by principals and teacher value-added measures. *American Educational Research Journal*, 51(1), 73–113.

- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. London: Routledge.
- Helsby, G. (1999). *Changing teachers' work: The reform of secondary schooling*. Philadelphia, PA.: Open University Press.
- Hill, D. (2007). Educational perversion and global neoliberalism. In E. W. Ross & R. J. Gibson (Eds.), *Neoliberalism and education reform* (pp. 107–144). Cresskill, NJ: Hampton Press.
- Honig, M. I. (2006). Complexity and policy implementation: Challenges and opportunities for the field. In M. I. Honig (Ed.), *New directions in education policy implementation: Confronting complexity* (pp. 1–23). Albany, NY: State University of New York Press.
- hooks, bell. (2003). *Teaching community: A pedagogy of hope*. New York, NY: Routledge.
- Hughes, S., Pennington, J. L., & Makris, S. (2012). Translating autoethnography across the AERA standards: Toward understanding autoethnographic scholarship as empirical research. *Educational Researcher*, 41(6), 209–219.
- Hursh, D. W. (2007). Marketing education: The rise of standardized testing, accountability, competition, and markets in public education. In E. W. Ross & R. J. Gibson (Eds.), *Neoliberalism and education reform*. Cresskill, NJ: Hampton Press.

- Jackson, C. K. (2012). Non-cognitive ability, test scores, and teacher quality: Evidence from 9th grade teachers in North Carolina. *National Bureau of Economic Research, Inc*, (December), 1–40.
- Jerald, C. D. (2009). *The value of value-added data. K-12 policy*. Washington, DC: Education Trust.
- Jiang, J. Y., Spote, S. E., & Luppescu, S. (2015). Teacher perspectives on evaluation reform: Chicago's REACH students. *Educational Researcher*, 44(2), 105–116.
<https://doi.org/10.3102/0013189X15575517>
- Johnson, S. M. (2015). Will VAMS reinforce the walls of the egg-crate school? *Educational Researcher*, 44(2), 117–126.
<https://doi.org/10.3102/0013189X15573351>
- Kelchtermans, G. (1996). Teacher vulnerability: Understanding its moral and political roots. *Cambridge Journal of Education*, 26(3), 307.
- Kelchtermans, G. (2005). Teachers' emotions in educational reforms: Self-understanding, vulnerable commitment and micropolitical literacy. *Teaching and Teacher Education*, 21(8), 995–1006. <https://doi.org/10.1016/j.tate.2005.06.009>
- Kelchtermans, G. (2009). Who I am in how I teach is the message: Self-understanding, vulnerability and reflection. *Teachers and Teaching*, 15(2), 257–272.
<https://doi.org/10.1080/13540600902875332>
- Kemmis, S. (2006). Participatory action research and the public sphere. *Educational Action Research*, 14(4), 459–476.

- Kim, J., & Youngs, P. (2016). Promoting instructional improvement or resistance? A comparative study of teachers' perceptions of teacher evaluation policy in Korea and the USA. *Compare: A Journal of Comparative and International Education*, 46(5), 723–744. <https://doi.org/10.1080/03057925.2015.1057478>
- Kincheloe, J. L. (1991). *Teachers as researchers : Qualitative inquiry as a path to empowerment*. London: Falmer Press. Retrieved from <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat00263a&AN=mts.b1334813&site=eds-live&scope=site>
- Lasky, S. (2005). A sociocultural approach to understanding teacher identity, agency and professional vulnerability in a context of secondary school reform. *Teaching and Teacher Education*, 21(8), 899–916. <https://doi.org/10.1016/j.tate.2005.06.003>
- Lee, L. (2011). What did the teachers think? Teachers' responses to the use of value-added modeling as a tool for evaluating teacher effectiveness. *Journal of Urban Learning, Teaching, and Research*, 7, 97–103.
- Lempert, L. B. (2007). Asking questions of the data: Memo writing in the grounded theory tradition. In K. Charmaz & A. Bryant (Eds.), *The SAGE handbook of grounded theory*. London, England: Sage Publications. Retrieved from <http://srmo.sagepub.com/view/the-sage-handbook-of-grounded-theory/SAGE.xml>
- Leyva, R. (2009). No Child Left Behind: A neoliberal repackaging of social Darwinism. *Journal for Critical Education Policy Studies*, 7(1), 364–381.
- Lezotte, L. W., & Snyder, K. M. (2010). *What effective schools do: Re-envisioning the correlates*. Bloomington, IN: Solution Tree.

- Lipman, P. (2007). "No Child Left Behind": Globalization, privatization, and the politics of inequality. In E. W. Ross & R. J. Gibson (Eds.), *Neoliberalism and education reform*. Cresskill, NJ: Hampton Press.
- Lomborg, K., & Kirkevold, M. (2003). *Truth and validity in grounded theory – a reconsidered realist interpretation of the criteria: fit, work, relevance and modifiability* (Vol. 4). <https://doi.org/10.1046/j.1466-769X.2003.00139.x>
- Lortie, D. C. (1975). *Schoolteacher: A sociological study*. Chicago, IL: University of Chicago Press. Retrieved from <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat00263a&AN=mts.b1064454&site=eds-live&scope=site>
- Madison, D. S. (2005). *Critical ethnography: Method, ethics, and performance*. Thousand Oaks, CA: SAGE Publications.
- Marshall, C. (1997). Dismantling and reconstructing policy analysis. In C. Marshall (Ed.), *Feminist critical policy analysis*. London: Falmer Press.
- Marshall, C., & Rossman, G. B. (2011). *Designing qualitative research* (5th ed.). Los Angeles: Sage Publications.
- Marx, K., & Engels, F. (1976). *Theses on Feuerbach: 3*. London: Lawrence & Wishart. Retrieved from <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edspmu&AN=edspmu.MUSE9781909831155.8&site=eds-live&scope=site>

- März, V., Kelchtermans, G., & Dumay, X. (2016). Stability and change of mentoring practices in a capricious policy environment: Opening the “Black Box of Institutionalization.” *American Journal of Education*, *122*(3), 303–336.
- Maxwell, J. A. (1996). *Qualitative research design: An interactive approach*. Thousand Oaks, Calif: Sage Publications.
- Maxwell, J. A. (2012). *Qualitative research design: An interactive approach*. Thousand Oaks, CA: Sage Publications.
- Mayo, P. (1999). *Gramsci, Freire, and adult education: Possibilities for transformative action*. New York, NY: Zed Books.
- Mayo, P. (2003). A rationale for a transformative approach to education. *Journal of Transformative Education*, *1*(1), 38–57.
- McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A., & Hamilton, L. (2004a). Let’s see more empirical studies on value-added modeling of teacher effects: A reply to Raudenbush, Rubin, Stuart and Zanutto, and Reckase. *Journal of Educational and Behavioral Statistics*, *29*(1), 139–143.
- McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A., & Hamilton, L. (2004b). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, *29*(1), 67–101. <https://doi.org/10.3102/10769986029001067>
- McLeroy, K. R., Bibeau, D., Steckler, A., & Glanz, K. (1988). An ecological perspective on health promotion programs. *Health Education & Behavior*, *15*(4), 351–377. <https://doi.org/10.1177/109019818801500401>

- McNeil, M. (2010, January 19). All but 10 states throw hats into Race to Top ring [Blog post]. Retrieved from https://www.edweek.org/ew/articles/2010/01/20/19rtt_ep.h29.html
- Merriam-Webster. (n.d.). Free market. *Merriam-Webster*. Retrieved from [https://www.merriam-webster.com/dictionary/free market](https://www.merriam-webster.com/dictionary/free%20market)
- Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340–363.
- Michalek, C. E. (2014). *A study of assessment results, teacher effectiveness ratings, and teacher perceptions under a value-added teacher evaluation model* (Doctoral dissertation). Retrieved from <http://search.proquest.com.ezproxy.mtsu.edu/pqdtft/docview/1666390644/abstract/89541F07C4844764PQ/7?accountid=4886>
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage Publications.
- Mitra, R. (2012). “My country’s future”: A culture-centered interrogation of corporate social responsibility in India. *Journal of Business Ethics*, 106(2), 131–147.
- Morse, J. M. (2004). Theoretical saturation. In M. Lewis-Beck, A. Bryman, & T. Futing Liao (Eds.), *The SAGE Encyclopedia of Social Science Research Methods*. Thousand Oaks, CA: Sage Publications. Retrieved from <http://methods.sagepub.com/reference/the-sage-encyclopedia-of-social-science-research-methods>

- Moustakas, C. (1994). Phenomenological research methods. In Norman K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 220–235). Thousand Oaks, CA: SAGE Publications.
- Mutua, K., & Swadener, B. B. (Eds.). (2004). *Decolonizing research in cross-cultural contexts: Critical personal narratives*. Albany, NY: State University of New York Press.
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform. A report to the Nation and the Secretary of Education, United States Department of Education* (No. 0455-B-2) (p. 84). Washington, DC.
- National Council on Teacher Quality. (n.d.). Measures of student growth, 2017 teacher and principal evaluation policy. Retrieved from <https://www.nctq.org/yearbook/national/Measures-of-Student-Growth-77>
- National Research Council. (2009). *Letter report to the U.S. Department of Education on the Race to the Top Fund*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/12780>
- Nichols, S. L., & Berliner, D. C. (2007). *Collateral damage: How high-stakes testing corrupts America's schools*. Cambridge, MA: Harvard Education Press.
- Noddings, N. (1992). *The challenge to care in schools : An alternative approach to education*. New York, NY: Teachers College Press.
- Orlitzky, M. (2016). Free market. *Encyclopaedia Britannica*. Retrieved from <https://www.britannica.com/topic/free-market>

- Palmer, P. J. (1998). *The courage to teach: Exploring the inner landscape of a teacher's life*. San Francisco, CA: Jossey-Bass.
- Papay, J. P. (2011). Different tests, different answers: The stability of teacher value-added estimates across outcome measures. *American Educational Research Journal*, 48(1), 163–193. <https://doi.org/10.3102/0002831210362589>
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.
- Patton, M. Q. (2005). Qualitative Research. In *Encyclopedia of Statistics in Behavioral Science*. Chichester, UK: John Wiley & Sons, Ltd. Retrieved from <http://dx.doi.org/10.1002/0470013192.bsa514>
- Patton, M. Q. (2015). *Qualitative research and evaluation methods: Integrating theory and practice* (4th ed.). Thousand Oaks, CA: Sage Publications.
- Piro, J., Wiemers, R., & Shutt, T. (2011). Using student achievement data in teacher and principal evaluations: A policy study. *International Journal of Educational Leadership Preparation*, 6(4), 1–9.
- Polkinghorne, D. E. (2007). Validity issues in narrative research. *Qualitative Inquiry*, 13(4), 471–486.
- Polkinghorne, D. E. (2010). The practice of narrative. *Narrative Inquiry*, 20(2), 392–396.
- Powell, W. W., & DiMaggio, P. (1991). *The new institutionalism in organizational analysis*. Chicago, IL: University of Chicago Press. Retrieved from <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cab00263a&AN=mts.b1336694&site=eds-live&scope=site>

- Prussia, G. E., & Kinicki, A. J. (1996). A motivational investigation of group effectiveness using social-cognitive theory. *Journal of Applied Psychology, 81*(2), 187–198.
- Rainwater, K. A. (2016, April 27). Tennessee cancels TNReady testing for grades 3-8, ends contract with test vendor. *Times Free Press*. Retrieved from <https://www.timesfreepress.com/news/politics/state/story/2016/apr/27/tennessee-terminates-contract-tnready-test-company/362575/>
- RAND Corporation. (2002). Value-added modeling 101: Using student test scores to help measure teaching effectiveness. Retrieved from https://www.rand.org/pubs/corporate_pubs/CP693z4-2012-09.html.
- Ratner, C. (2000). Agency and culture. *Journal for the Theory of Social Behaviour, 30*(4), 413.
- Raver, C. C., Blair, C., & Willoughby, M. (2013). Poverty as a predictor of 4-year-olds' executive function: New perspectives on models of differential susceptibility. *Developmental Psychology, 49*(2), 292–304. <https://doi.org/10.1037/a0028343>
- Reed-Danahay, D. (2005). *Locating Bourdieu*. Bloomington, IN: Indiana University Press. Retrieved from <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat00263a&AN=mts.b3015925&site=eds-live&scope=site>
- Reeves, D. B. (2005). A constructive alternative in a destructive debate. *Principal Leadership, 5*(7), 38–43.

- Reio Jr., T. G. (2005). Emotions as a lens to explore teacher identity and change: A commentary. *Teaching and Teacher Education, 21*(8), 985–993.
<https://doi.org/10.1016/j.tate.2005.06.008>
- Rigby, J. G., Woulfin, S. L., & Marz, V. (2016). Understanding how structure and agency influence education policy implementation and organizational change. *American Journal of Education, 122*(3), 295–302.
- Risley, T. R., & Hart, B. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Paul H. Brookes Publishing Co.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417–458.
- Roulston, K. (2010). *Reflective interviewing: A guide to theory and practice*. Los Angeles, CA: Sage Publications.
- Ruggunan, S., & Spiller, D. (2014). Critical pedagogy for teaching HRM in the context of social change. *African Journal of Business Ethics, 8*(1), 29–43.
- Saldaña, J. (2012). *The coding manual for qualitative researchers* (2nd ed). Los Angeles, CA: Sage Publications.
- Sallis, J. F., & Owen, N. (2015). Ecological models of health behavior. In K. Glanz, B. Rimer, & K. Viswanath (Eds.), *Health behavior and health education: Theory, research, and practice* (5th ed., pp. 43–64). San Francisco, CA: Jossey-Base.
- Sanders, W. L., & Horn, S. P. (1998). Research findings from the Tennessee Value-Added Assessment System (TVAAS) database: Implications for educational

evaluation and research. *Journal of Personnel Evaluation in Education*, 12(3), 247.

Sanders, William L., & Horn, S. P. (1994). The Tennessee Value-Added Assessment System (TVAAS): Mixed-model methodology in educational assessment. *Journal of Personnel Evaluation in Education*, 8(3), 299–311.

SAS Institute. (2017). *Technical documentation of 2017 TVAAS evaluation*. Retrieved from <https://tvaas.sas.com/support/TVAAS-TechnicalDocumentation-2017.pdf>

Schmidt, M., & Datnow, A. (2005). Teachers' sense-making about comprehensive school reform: The influence of emotions. *Teaching and Teacher Education*, 21(8), 949–965. <https://doi.org/10.1016/j.tate.2005.06.006>

Schochet, P. Z., & Chiang, H. S. (2013). What are error rates for classifying teacher and school performance using value-added models? *Journal of Educational and Behavioral Statistics*, 38(2), 142–171.

Schwandt, T. A. (1998). Constructivist, interpretivist approaches to human inquiry. In Y. S. Lincoln & N. K. Denzin (Eds.), *The landscape of qualitative research theories and issues* (pp. 221–259). Thousand Oaks, CA: Sage Publications.

Schwandt, T. A. (2007). *The Sage dictionary of qualitative inquiry*. Los Angeles, CA: Sage Publications.

Scott, W. R. (2005). Institutional theory. In G. Ritzer (Ed.), *Encyclopedia of social theory* (Vol. 11, pp. 408–414). Thousand Oaks, CA: Sage Publications. Retrieved from <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=e000xna&AN=474409&site=ehost-live&scope=site>

- Scott, W. R. (2014). *Institutions and organizations: Ideas, interests, and identities* (Fourth edition). Los Angeles: SAGE Publications.
- Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information, 22*(2), 63–75.
- Smith, A. (1776). *The Wealth of Nations*.
- Spillane, J. P., & Burch, P. (2006). The institutional environment and instructional practice: Changing patterns of guidance and control in public education. In H.-D. Meyer & B. Rowan (Eds.), *The New Institutionalism in Education* (pp. 87–102). Albany, NY: State University of New York Press. Retrieved from <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=e000xna&AN=179765&site=ehost-live&scope=site>
- Stokols, D. (1992). Establishing and maintaining healthy environments: Toward a social ecology of health promotion. *American Psychologist, 47*(1), 6.
- Stronge, J. H. (2007). *Qualities of effective teachers* (2nd ed). Alexandria, VA: Association for Supervision and Curriculum Development.
- Tàbara, J. D., & Chabay, I. (2013). Coupling Human Information and Knowledge Systems with social–ecological systems change: Reframing research, education, and policy for sustainability. *Environmental Science & Policy, 28*, 71–81. <https://doi.org/10.1016/j.envsci.2012.11.005>
- Tennessee Department of Education. (2012). *Teacher evaluation in Tennessee: A report on year 1 implementation*.

- Tennessee Department of Education. (n.d.). Tennessee Value-Added Assessment Aystem. Retrieved from <https://www.tn.gov/education/data/tvaas.html>
- The Tennessean. (2017, March 18). Dr. William “Bill” Sanders. Retrieved from <https://obits.tennessean.com/obituaries/tennessean/obituary.aspx?n=william-sanders-bill&pid=184563925&fhid=14192>
- Thomas, W. I., & Thomas, D. S. (1928). *The child in America: Behavior problems and programs*. New York, NY: Alfred A Knopf.
- Thornton, P. H., & Ocasio, W. (2008). Institutional logics. In *The SAGE handbook of organizational institutionalism* (pp. 99–129). Retrieved from <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edszbw&AN=EDSZBW592503364&site=eds-live&scope=site>
- Tienken, C. H. (2013). Neoliberalism, social Darwinism, and consumerism masquerading as school reform. *Interchange: A Quarterly Review of Education*, 43(4), 295–316.
- Tucker, K. H. (1998). *Anthony Giddens and modern social theory*. London: Sage Publications. Retrieved from <https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=309851&site=eds-live&scope=site>
- Tucker, M. S. (2014). *Fixing our national accountability system*. Washington, DC: National Center on Education and the Economy.
- Turner, R. S. (2008). *Neo-liberal ideology: History, concepts and policies*. Edinburgh, Scotland: Edinburgh University Press. Retrieved from

<https://ezproxy.mtsu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cat00263a&AN=mts.b3028841&site=eds-live&scope=site>

- Tuytens, M., & Devos, G. (2009). Teachers' perception of the new teacher evaluation policy: A validity study of the Policy Characteristics Scale. *Teaching and Teacher Education, 25*(6), 924–930. <https://doi.org/10.1016/j.tate.2009.02.014>
- Tuytens, M., & Devos, G. (2014). The problematic implementation of teacher evaluation policy: School failure or governmental pitfall? *Educational Management Administration & Leadership, 42*(4), 155–174.
- Uitto, M., Kaunisto, S.-L., Kelchtermans, G., & Estola, E. (2016). Peer group as a meeting place: Reconstructions of teachers' self-understanding and the presence of vulnerability. *International Journal of Educational Research, 75*, 7–16. <https://doi.org/10.1016/j.ijer.2015.10.004>
- Urquhart, C. (2013). *Grounded theory for qualitative research: A practical guide*. Los Angeles, CA: Sage Publications. Retrieved from <http://dx.doi.org.ezproxy.mtsu.edu/10.4135/9781526402196>
- U.S. Congress. No Child Left Behind Act of 2001, Pub. L. No. 107–110 (2002). Retrieved from <https://www.gpo.gov/fdsys/pkg/PLAW-107publ110/html/PLAW-107publ110.htm>.
- U.S. Department of Education. (2009). *Race to the Top Executive Summary*. Retrieved from <https://www2.ed.gov/programs/racetothetop/executive-summary.pdf>

- U.S. Department of Education. (2010, March 29). Delaware and Tennessee win first Race to The Top grants. Retrieved from <https://www.ed.gov/news/press-releases/delaware-and-tennessee-win-first-race-top-grants>
- van den Berg, R., Vandenberghe, R., & Sleegers, P. (1999). Management of innovations from a cultural-individual perspective. *School Effectiveness and School Improvement, 10*(3), 321–351.
- Van Manen, M. (1990). *Researching lived experience: Human science for an action sensitive pedagogy*. New York, NY: State University of New York Press.
- van Veen, K., Sleegers, P., & van de Ven, P.-H. (2005). One teacher's identity, emotions, and commitment to change: A case study into the cognitive–affective processes of a secondary school teacher in the context of reforms. *Teaching and Teacher Education, 21*(8), 917–934.
- Wadsworth, Y. (2011). *Everyday evaluation on the run*. Walnut Creek, CA: Left Coast Press.
- Welsch, D. M., & Zimmer, D. M. (2010). The effect of health and poverty on early childhood cognitive development. *Atlantic Economic Journal, 38*(1), 37–49.
- Wood, R., & Bandura, A. (1989). Social cognitive theory of organizational management. *Academy of Management Review, 14*(3), 361–384.
<https://doi.org/10.5465/AMR.1989.4279067>
- Woods, P., & Jeffrey, B. (2002). The reconstruction of primary teachers' identities. *British Journal of Sociology of Education, 23*(1), 89–106.

Woulfin, S. L. (2016). Duet or duel? A portrait of two logics of reading instruction in an urban school district. *American Journal of Education, 122*(3), 337–365.

Zembylas, M., & Barker, H. B. (2007). Teachers' spaces for coping with change in the context of a reform effort. *Journal of Educational Change, 8*(3), 235–256.

APPENDICES

APPENDIX A
TABLE OF CODES

Core Themes	Categories from Question Analysis	Open Codes
Teachers distrust value-added methods and evaluation.	Questionable validity of value-added models VAM	Fairness, Differences Across Grade Levels, Doesn't Measure Learning, Outside Factors,
	Problems with application of policy	Timing of Scores, No Feedback,
	Issues with TCAP	Anxiety, Effect on Students, Not Developmentally Appropriate, Pressure
	Uncertainty and lack of understanding of VAM	Secretive, Unclear, Lack of Understanding
	Negative effect on teaching	Lose Love of Teaching, Lose the Art
	Effect on relationships with colleagues	Increased Competition, Stress
	Inter-school competition	Competitiveness, Lack of Communication, Ranking
Institutional practices establish norms in teacher evaluation.	Strategic Teacher Placement	Improve Scores, Teacher Fit
	Data Day (Purpose)	Reflection, Motivator, Analyze Scores, Accountability, Public, Teacher Placement
	Data Response Team (DRT)	Analyze, Public, District Data Person
	Additional Resources	Response to Intervention, Teaching Assistants, Time
	Testing Support	Student Meals, Testing Schedule
	Efficacy-oriented interactions with principal	School Has a Stake, Leadership Opportunities, Earned Freedom
	Action-oriented interactions with principal	Increase Growth, Close to the Line, Focus for Growth, Motivation, Model Teaching
	Data-oriented interactions with principal	Discuss Scores

Table of Codes cont.

Core Themes	Categories from Question Analysis	Open Codes
Institutional practices and policies impact teacher self-esteem.	Data Day (Teacher Self-Esteem)	Shame, Embarrassed, Upset, Pressure, Failure, Humiliating, Devastating, Horrible, Crying
	Efficacy-oriented interactions with principal	Praise, Encouragement, Stress, Embarrassing, Questioned, Scores Not Used Right
	Negative opinion of TVAAS	Failure, Leave Profession, Pressure, Shaming, Stress, Teacher as a Number, Low School Morale
Teachers and teacher teams choose practices to improve student scores.	Test Preparation	Content, Format, Parent Support, Practice Tests, Review, Stamina, Testing Strategies, Testing Tools
	Planning	Content, Looking Back, Standards, Show Growth
	Instruction	Individualize, Small Groups
	Curriculum	Pacing, Standards
	Departmentalization	Student Growth, Move Kids, Better Scores, Expert, Focus, Standards, Relief, Volume of Standards, Mastery, Teacher Fit, Positive, Excited, Better Teacher
	Classes grouped by student ability	Student Growth, Improved Scores, Differentiation, Scaffolding, Small Groups, Move Students Faster, Move More Students, Success
	Informal interactions	Share Strategies, Different Options, Supportive, Test Prep, Find Gaps
	Formal interactions	Curriculum, (Vertical) Planning, Review (TVAAS) Scores, Share Strategies, Extended Instruction

APPENDIX B

Teacher interview guide for grounded study, Teachers "talk back": Exploring the dynamics between practice and value-added evaluation.	
Making Sense of Institutional Logics	
<i>Personal Interpretive Framework</i>	
Professional History	
<input type="checkbox"/>	How did you become a teacher?
<input type="checkbox"/>	How would you describe yourself as a teacher?
Professional Present	
<input type="checkbox"/>	Please tell me about your beliefs about good teaching.
<input type="checkbox"/>	Please tell me about the expectations you hold for yourself as a teacher
<i>Beliefs About Teacher Evaluation</i>	
<input type="checkbox"/>	In your opinion, what is the best use of teacher evaluations?
<i>Logics of Value-Added Scores in Evaluation Policy</i>	
<input type="checkbox"/>	In your opinion, what is the intended purpose of the value-added component in teacher evaluations (as part of the Tennessee Educator Acceleration Model (TEAM))?
<input type="checkbox"/>	What experiences lead you to this understanding of value-added evaluations?
Practices Related to Value-Added Evaluation Policy	
<i>Impact on Practice</i>	
<input type="checkbox"/>	Does the way you're evaluated shape your teaching? Why or why not?
<input type="checkbox"/>	If yes; please tell me about how your anticipated value-added score(s) shape(s) your teaching practices. (content, methods)
<i>Impact on Interactions</i>	
<input type="checkbox"/>	Please describe interactions you have had with colleagues about value-added scores.
<input type="checkbox"/>	Have any of these interactions had an impact on your teaching? And if yes, how?
<input type="checkbox"/>	Please describe any interactions you have had with principal about value-added scores.
<input type="checkbox"/>	Have your interactions with your principal about value-added scores had an impact on your teaching? If yes, how?
<i>School-level Activities and Structures Related to Teacher Evaluation</i>	
<input type="checkbox"/>	Please describe the ways, if any, that your team adapts or organizes instruction in anticipation of value-added scores.
<input type="checkbox"/>	Please describe the ways you are aware of, if any, that your principal has implemented team or school-wide practices related to your school's value-added scores.
Closing Questions	
<input type="checkbox"/>	Is there something that occurred to you during our interview that you would like to share?
<input type="checkbox"/>	Is there something else you think should know to understand your thoughts and experiences?
<input type="checkbox"/>	Is there anything you would like to ask me?