THE UNIVERSITY OF CHICAGO

EXCELLENCE FOR ALL: PURSUING EFFECTIVENESS IN A HIGH-PERFORMING, NO-EXCUSES CHARTER SCHOOL

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF SOCIOLOGY

BY

JASON RADFORD

CHICAGO, ILLINOIS

DECEMBER 2019

For my father Lonnie who was there for the beginning but not the end.

If the central purpose of education research is to identify solutions and provide options for policymakers and practitioners, one would have to characterize the past five decades as a near-complete failure. (Kane 2016)

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I first need to thank the teachers, staff, and administrators at Achievement College Prep who welcomed me into their school, their community, their private spaces, their one-on-one meetings, and bitch fests. It takes a lot of courage to license one's life and work to someone you barely know with no guarantee anything will come of it or that the result will be flattering. I especially want to thank the school's executive director, the dean of instruction, and the chairs of the math and English departments who provided me with the most access to their work, warts and all.

I would also like to thank my own teachers, starting with Dr. Elisabeth Clemens, who shepherded me through ten years of Chicago sociology starting with my first year in the Master of Arts Program in the Social Science. Lis believed in me before there was good reason to and has guided me through all of the twists and turns and ups and downs of graduate school and becoming an academic. I'll miss coming to her office to glean her preternatural ability to find solutions in my own work and to frame it in new and enlightening ways, pointing the ship ever forward to its destination.

The members of my committee were there when I needed them. Ron Burt, ever positive, supportive, generous, and direct; was a bottomless source of strength, insight, and story. John Levi Martin always started kooky, but inevitably opened whole new ways of seeing my work. Forrest Stuart helped me find my footing in my fieldwork, molding the raw, ethnographic clay given me by Omar McRoberts. I will miss them all dearly.

I must also thank the faculty who left their indelible mark on the life of this mind. First and foremost is James Evans who helped me start the Computational Social Sciences workshop and continues to help me push the boundaries of social science method. To Karin Knorr whose

work and teaching on the social construction of science always finds its way to the heart of my research. And, to Kristen Schilt who I'll never learn enough from.

Completing the dissertation is not only a feat of research, but of administration. I would not be here without the help of Linnea Martin and Pat Princell helping me navigate the university and ensuring I had what I needed.

I shall also thank the many friends who offered encouragement and support along the way: Betsy Sinclair, John Padgett, and John Brehm all offered warmth, support, and intellectually engaging conversation. My fellow graduate students: Peter McMahan and Michael Castelle who taught me to code, my roommate Jim Murphy, Misha Teplitskiy, Katie Hendricks, Jeffrey Parker, Peter Fugiel, and Nick Judd.

Special thanks are deserved for David Lazer and my friends at Northeastern University who offered encouragement and commiseration as I worked on my dissertation.

Finally, I would like to thank my family starting with my wife Rachel who waited in patience and in poverty. My mom and dad always supported me unconditionally, even as I took a nontraditional path and didn't always make the best decisions with the forbearance they gave.

CHAPTER 1

EDUCATION IN THE WILDERNESS

On December 10, 2015, President Barack Obama put an end to the consensus at the heart of education reform. For the past thirty-five years, education reform in America had been focused on setting high expectations for students, providing high quality teachers and schools, and ensuring all students from all backgrounds and abilities get a high-quality education. In his bill-signing speech, Obama declared victory saying,

> We've seen states raise academic expectations for all students. That means that we're in a better position to out-teach and out-compete other nations at a time when knowledge is really the single-biggest determinant of economic performance. High school graduation rates have reached an all-time high; dropout rates have hit historic lows. The number of high schools so bad they're called "dropout factories" has been cut almost in half. We're training tens of thousands of outstanding math and science teachers. More students are graduating from college than ever before, and more than a million additional black and Hispanic students are now going to college.

Despite the plaudits, the bill itself unwound the federal education policy that had been articulated in 1983 and written into law by the No Child Left Behind Act in 2002. It sought to reduce standardized testing, decentralized learning standards and school assessment, and shifted research funding to nonprofit and local community organizations. In other words, the federal government was backing away from making schools more effective, leaving it to the states to figure out.

THE ERA OF EXCELLENCE

The end of the era has gone unnoticed. It began with the publication of the influential report *A Nation at Risk*. The titular risk was that we were losing our economic competitiveness to countries like Japan, Korea, and Germany whose children were learning more than ours.

According to data at the time, American students in 1983 were learning less than students in 1963. The report's authors ascribed this erosion of American learning to "a rising tide of mediocrity" citing shorter school days, frivolous coursework, and poorly-managed classrooms. , Compared to their Asian and European counterparts, American students were spending less time in school and less time on core, rigorous courses led by qualified teachers. The report's authors stressed the need to create an education system that provided an excellent education for everyone rather than pumping out mediocre students.

Their recommendations for achieving this formed the foundation of education reform as most in the U.S. know it today: require more coursework in the "Core Four" subjects of science, math, English, and social studies; use standardized tests to control what students get promoted to the next grade; hire better teachers and set higher standards for them; ensure minority, disabled, and disadvantaged students are treated equally; and study what policies, programs, and initiatives work in the pursuit of these goals. Rather than minting new graduates, American schools were to provide excellent education to all.

In the subsequent twenty years of this era of excellence, states across the country adopted the point-of-view and policies outlined in *A Nation at Risk*. States implemented core course requirements, created curricular standards for those courses, and made standardized tests a prerequisite for student advancement. They created new, alternative pathways for technical professionals to enter teaching. They began publishing data and creating school report cards by which to evaluate their performance.

A decade after the report, a different movement for excellence in education began to take root. This movement sought to create whole new schools and school systems devoted to improving education. These so-called charter schools would be labs of experimentation, meant to

find the best ways to teach students that were otherwise disengaged, hard to reach, or facing unique challenges. They would be given extra freedom from laws, policies, and rules in order to develop innovative practices. By 2000, more than three quarters of states had passed laws authorizing the creation of charter schools.

The consensus around educational excellence reached its apex in the No Child Left Behind Act (NCLB). The act created a federal mandate that students hit minimum learning outcomes called "Adequate Yearly Progress" (AYP) or else schools would risk lose funding. The act was designed to force all schools to achieve excellence or perish. Yet, NCLB still relied on states to define and measure AYP. In 2010, the Obama Administration ratcheted standards and testing up through its Race to the Top program which encouraged all but a handful of states to adopt the Common Core.

What is amazing about this era is its focus and bipartisanship. There was a substantial degree of consensus in both parties that excellence was the proper goal of public education. The NCLB passed overwhelming in both chambers of congress. Republican- and democrat-controlled states implemented these same basic policies. And yet, it has all come unraveled in only a couple years. Obama's retrenchment, the Every Student Succeeds Act, passed with 85% of the vote in congress in 2015. It is not that we have given up on excellence or the belief that testing, standards, and high expectations will get us there. But the past thirty-five years of policy have amounted to very little in the way of real improvements in how much American students learn.

SO MUCH REFORM, SO LITTLE EFFECT

A Nation at Risk established a pinnacle of ambition for public education in America. Every student, no matter the conditions of their birth whether race, ethnicity, income, or ability would receive a great education in order to ensure social equality and the economic competitiveness of the country. This may seem to have been the aspiration of schools all along, but it was never the practice. Whether segregation, high dropout rates, or tracking students into technical professions; America long maintained a system of education that allowed substantial numbers of students to not take academic classes, not graduate, or to not attend school altogether. The notion that we would create an education system in which all schools would be good schools and that all students, no matter their learning challenges, would receive a high quality education represented a sea change in governmental aspirations.

What policymakers meant for a school to be good and for students to have received a good education boils down to standardized test scores. Good schools are where all students score well on tests and these scores represent students' learning. This impetus to focus on test scores for students from all backgrounds has become the language of closing various gaps in education. These gaps are the heart of ascriptive inequality – measuring the disadvantage bestowed to America's children by the inalterable conditions of their birth. The central policy problem of our time is creating an educational system capable where students from all groups achieving full proficiency.

Paralleling our increasingly ambitious vision for education has been a steady growth in spending on education and a bevy of research on how to improve student learning. Figure 1.1 shows the growth in government spending per pupil in education throughout the twentieth

century. While there are troughs in the early 1980s and throughout the 1990s, the trend is upward until the financial crisis in 2010.

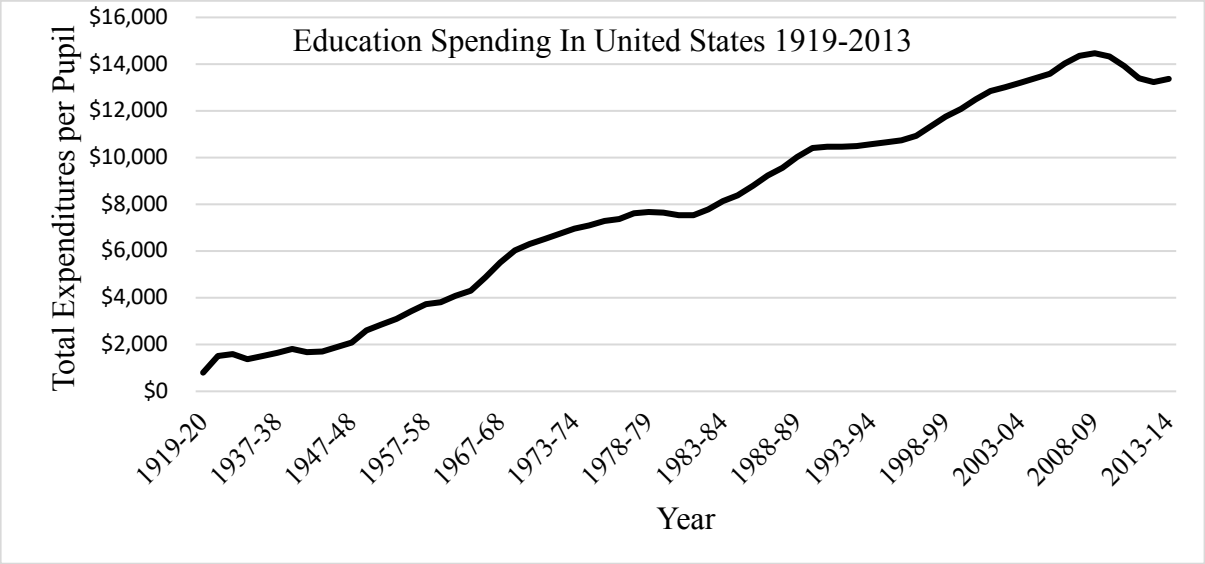

Figure 1.1 Education Spending in the United States 1919-2013.
NOTE: Total expenditure is per pupil in daily attendance in 2015-2016 dollars.
SOURCE: U.S. Department of Education, National Center for Education Statistics, Digest of Educational Statistics, Table 236.55.

The demand for better research was met. The push to standardized testing and school accountability led to high quality data on student, teacher, and school performance in states like North Carolina, Tennessee, Florida, and Ohio. Researchers have taken advantage of the use of lottery systems which randomize students to schools, to estimate the causal effect of schools and school policy on learning (Angrist, Bettinger, and Kremer 2006; Angrist et al. 2017; Clark et al. 2015; Cullen, Jacob, and Levitt 2006; Dobbie and Fryer 2013). Every year our knowledge of what matters and how much to a students' education improves.

As Thomas Kane (2014), the prominent economist of education, has written:

> In 25 years as an education researcher, I have never witnessed a rapid outpouring of new research in education such as we've seen in recent years. In what may be the most important byproduct of the No Child Left Behind Act and the Institute of Education Science's grants to states, school agencies have been linking students to teachers and

schools and tracking their achievement over time. Researchers across the country have been using those data to study the value of traditional teacher certification, the degree of on-the-job learning among teachers, the impact of charter schools, the effectiveness of teacher preparation programs, etc.

Finally, technology has improved our ability to deliver content to more students more effectively. When the NCLB was signed into law, 9% of households in the U.S. had high speed internet. Today, three in four households do (Pew Research Center 2019). Schools can now offer any number of advanced, collegiate courses online. From 2003 to 2010, the proportion of schools offering such distance education has increased 50% from 36% to 55% (National Center for Education Statistics 2012). The opportunities for students to learn and the efficiency with which we can deliver advance material are unprecedented.

These incredible gains in money, knowledge, and technological innovation and the rare consensus on what public education should provide should have revolutionized learning. And, yet, their impact appears to have been minimal. The best measures we have, two tests that have remained fairly constant over time and taken by students across many countries, show minimal growth in student learning during this period. The National Assessment of Education Progress (NAEP), called the "Nation's Report Card" and the Program for International Student Assessment (PISA), have been measuring students learning in a consistent and comparable way over time. These tests enable us to compare how much our schools are teaching students' over time.[1]

[1] Like all standardized tests, PISA and NAEP are not objective measures of learning. They select a certain set of skills and knowledge to test often without regard to the pedagogical priorities of individual states or countries. They should not be considered comprehensive or unbiased.

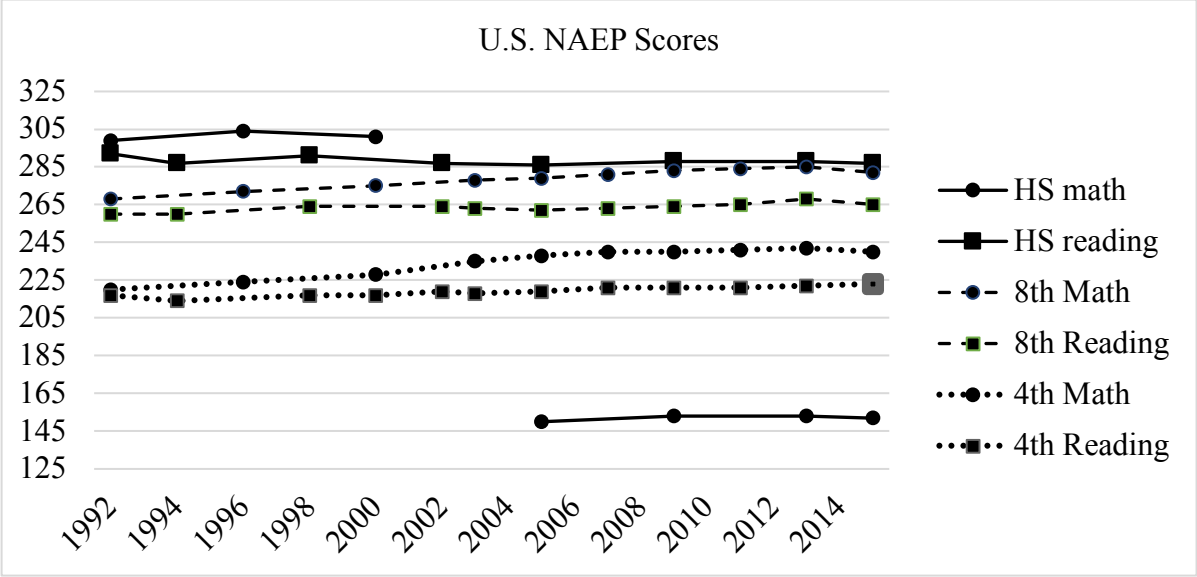

Figure 1.2 NAEP Scores 1992-2015.
SOURCE: NAEP Data Explorer, National Center for Education Statistics https://nces.ed.gov/nationsreportcard/naepdata/dataset.aspx.

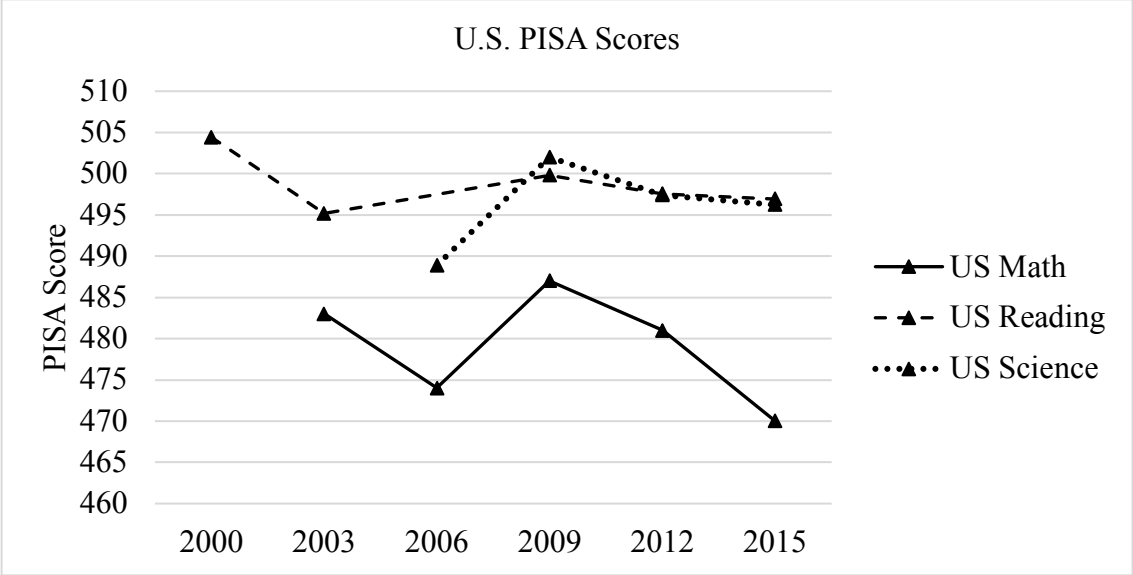

Figure 1.3 U.S. PISA Scores 2000-2015.
SOURCE: International Data Explorer, National Center For Education Statistics, http://nces.ed.gov/surveys/pisa/idepisa https://nces.ed.gov/timss/idetimss/dataset.aspx.
NOTE: the PISA tests students at the age of 15.

Figure 1.2 and Figure 1.3 show the results of the NAEP and PISA standardized tests for students in the United States since 1990. Over the past twenty years, we have seen an increase in two of the six core tests. These gains have been relatively small: 9% growth for fourth grade math and 7% for eighth grade math. Tellingly, there has been no change in high school grades. If school were providing more education, we should see the greatest amount of growth by the end of it rather than at its beginning.

In these same two decades, students, families, communities, and the nation have changed. The students who took the NAEP in 1992 are different from the students who took it in 2015. Students today are more likely to graduate from high school, more likely to be English Language Learners, be raised in single-parent homes, and more likely to have disabilities. Increasing educational effectiveness may be statistically counter-balanced by an increasingly difficult to teach population of students.

However, declining rates of teen fertility, crime, mortality, and lead poisoning during the same period have removed substantial barriers to education. Further, the increasing rates of education by previous generations mean that the children of the newly educated are more prepared to succeed than their parents were. Thus, schools may in fact be less effective than they used to be.

It is difficult to say whether students are more or less prepared today net these social changes. What is clear is that despite political support and considerable economic investment, all this talk and effort has produced little in the way of overall improvement in student achievement according to our best metrics. This dissertation asks why.

THE FLAW IN OUR MODEL

After a century of research on education, we think we know how to teach students well. Schools need more funding and smaller class sizes (C. K. Jackson, Johnson, and Persico 2016). They need strong school culture (A. S. Bryk and Schneider 2003; Sebring et al. 2006; Thapa et al. 2013). Teachers shouldn't be talking at students all day but getting them engaged in learning activities (Freeman et al. 2014). We think we know these to be true, but our track record over the past thirty years in creating effective schools, teachers, and policies demonstrates otherwise.

Creating Effective Schools

One of the central policy tools invented during the era of excellence was the school turnaround. Made popular by No Child Left Behind, school turnarounds represent the culmination of an emerging set of practices variously called "comprehensive school reform," "school reconstitution," and "school takeovers" started in the 1980s (Rice and Malen 2003; Wong and Shen 2003). Conceptually, turnaround has meant using government interventions to turn low-performing schools into high-performing ones. Practically, these interventions have ranged from bringing in consultants, changing teachers and staff, closing the school, or turning the school into a charter school.

Evaluations of school turnaround policies and programs are scarce, but those that exist indicate the results are mediocre or worse (Peck and Reitzug 2014). Aladjem et al. (2010) found only 10% of turnaround schools made consistent gains in student achievement within five years. Stuit (2010) finds that less than 2% of turnaround schools sustained any improvements after the first year of a turnaround intervention. Thompson et al. (2011) found that student achievement for turnaround schools in North Carolina did rise among high schools but not middle schools.

Broader research on school improvement indicates mediocre evidence for school reform programs in general. In their reviews, Herman et al. (1999) and Borman et al. (2003) found that only 3 of 29 school reform efforts demonstrated statistically meaningful effects. Aladjem et al. (2006) report only one had a measurable effect. While most programs were ineffective, those that were effective had an average effect that was larger than previous approaches to school reform. Unfortunately, they were still mild in relation to the overall distribution of schools (one-eighth of a standard deviation).

Overall, there is no evidence that governments can make schools better through turnaround interventions. There is very little research in general on what programs are effective at improving school performance. And this limited work suggests we have created one or two programs that produce marginal improvements in schools. For all the data we've created, policy focus on school improvement, and money on research; we have few proven options to turn to when it comes to improving schools. The same is true for teachers.

Creating Effective Teachers

If the evidence for our ability as a society to create effective schools is weak, the evidence for creating effective teachers is grim (Hill, Beisiegel, and Jacob 2013). First, human capital seems to matter little. Years of education, obtaining a master's degree, performance in college, and the prestige of a teacher's school of origin do not predict whether a teacher will be effective in the classroom (Aaronson, Barrow, and Sander 2007; Betts, Zau, and Rice 2003; Clotfelter, Ladd, and Vigdor 2006; Kane, Rockoff, and Staiger 2008). It is certainly doubtful that anyone off the street could teach. But once teachers have made it through the standard collegiate credentialing process, there is little impact to more or better education.

Once on the job, most teachers tend to go through a stable trajectory (see Rice 2010 for a review). They learn the most in the first two years on the job, performing poorest during their first year. From year two to year five, teachers typically make gradual improvements in their teaching. After year five, teachers plateau, becoming no more (or less) effective no matter how long they remain in the classroom. This plateau occurs regardless of how effective the teacher is when they plateau. There is some evidence that teachers in high school improve with experience, but it is weak and inconsistent (Harris and Sass 2011, see Clotfelter, Ladd, and Vigdor (2010) for an exception).

Finally, as with school turnaround, research on the efficacy of professional development programs is woefully sparse and variable (Hill, Beisiegel, and Jacob 2013). This is particularly shocking because teachers spend more hours in professional development each year than nearly every other profession including medical doctors.[2] Observationally, TNTP (formerly The New Teacher Project) found that the number of hours in professional development does not correlate with teacher effectiveness (2015).

In a meta-analysis of professional development programs in math, Gersten et al (2014) only found five studies meeting evidentiary standards for estimating effectiveness and only two of the five studies reported a positive result. In their review, Yoon et al (2007) find nine studies providing quality data from 1980 to 2005. These studies suggested receiving substantial professional development could increase student learning up to 21%. Thus, there is some evidence we could increase teachers' effectiveness with professional development, but the majority of research indicates we are not.

[2] Teachers – 155 (TNTP 2015), Doctors – 50 (D. Davis et al. 1999),

Creating Effective Policies

High-stakes testing is the hallmark of the era of excellence. Developed in the 1980s and implemented across most states throughout the 1990s, high-stakes testing was meant to control student promotion and incentivize schools to improve their educational by rewarding or punishing schools based on their test results. Since the 1990s, student test scores on state tests have increased dramatically (Kober, Chudowsky, and Chudowsky 2008). However, most of these gains appear to have been a chimera.

The problem is that, though students' scores were going up on state tests, they were not going up on the PISA or NAEP. Even before the NCLB mandated high-stakes testing in 2002, the evidence was mixed for whether such policies led to gains in student achievement. Some studies found that students in states with high-stakes testing performed no better on NAEP than those in states without it (Amrein-Beardsley and Berliner 2002; Nichols 2007; Nichols, Glass, and Berliner 2006). Others have found that states with high-stakes testing had higher scores on NAEP than those that did not, though longitudinal analyses show smaller or null effects (Carnoy and Loeb 2002; Hanushek and Raymond 2005; Klein et al. 2000; Koretz and Barron 1998). Ironically, the passage of NCLB has made it more difficult to study the impact of high-stakes testing because there are no longer any control groups of schools or students that are not tested. However, longitudinal analyses indicate NCLB-mandated testing continued the trend of weak and mixed results (Dee and Jacob 2011; Lee and Orfield 2006; Reback, Rockoff, and Schwartz 2014).

In all, there is moderate evidence that high-stakes testing has small, but real effects for 4th and 8th grade math, but sparse evidence that there is any improvement in reading or at the high school level. This suggests that teachers are "teaching to the test" – what students are learning in

one state, does not translate into a general education (Jennings and Bearak 2014; Jennings and Lauen 2016).

The Leap from Explanation to Control

Decades of research on what makes teachers, schools, and school systems more or less effective has yet to translate into substantial, sustained gains in student learning or teacher and school effectiveness. The problem is that knowing what works gives us very limited purchase on what *will* work.

First of all, our approach to assessing the impact of teacher quality, school culture, or student behavior misleads us into believing that all you need to do is increase these variables to make schools more effective. Thinking of schools in this way is a reification of the linear model (Abbott 1988b), encouraging us to over-emphasize the stable attributes of schools like culture, per pupil spending, and number of English-language learners rather than the internal dynamics by which culture is achieved, money spent, and English taught. In contrast to the assumptions of the statistical models that dominate social science, these dynamics are not independent from one another and have many different causal patterns that change over time.

Experimental and quasi-experimental methods are not immune to these fallacies. In fact, they encourage a second grave leap in logic. Knowing what works provides no insight into how to *make* it work. There is a large body of evidence showing that students taught by effective teachers learn more than those taught by ineffective teachers. But this result gives us no purchase on how to identify effective teachers, let alone train them. The recent "credibility revolution" only adds to our false sense of control (Angrist 2004; Angrist and Pischke 2010; M. Jackson and Cox 2013).

Understanding the inputs of effective education is not enough. We do not need another study showing that good teachers teach students more than bad teachers. We generally understand the ingredients for making education work. As with cooking, having the right ingredients does not make you a good chef. You must understand the processes by which those ingredients are transformed into delicious meals. We need a science of education that accounts for how culture, good teachers, and the like are created and sustained.

A NEW HOPE?

We know what ingredients good schools need but have failed to create effective programs implementing them. There does appear to be an important exception: no-excuses charter schools. These schools, focused on educating marginalized and disadvantaged students through high expectations and hard-nosed discipline, appear to have been able to create more effective schools. For this dissertation, I want to know how they managed to do it by performing a case study of one such school which I pseudonymously call Achievement College Prep.

I chose to perform a case study of a single charter school focusing on four initiatives members of the school were pursuing to become more effective. I studied these four initiatives through multiple lenses, times scales, and levels of analysis in order to answer the following questions. What were the changes being pursued and why? How were they formulated and implemented? What were the results? This in-depth, multi-perspectival look at the process of pursuing excellence at one effective school provides a single, well-studied case of the ways in which the attempt to create an effective school succeeds (and fails).

The case study is one of the best ways to build theory where none exists (Eisenhardt 1989; Eisenhardt and Graebner 2007). By studying one organization, I enumerate a broad set of

mechanisms, processes, interactions, and outcomes which contribute to the school's success. Not all schools pursue effectiveness in the same way or face the same issues. The case study is limited in that way. However, the results of this study should provide a starting point for more research on how the ingredients of success are translated into effective education. And, it should provide insight into other no-excuses schools facing similar challenges.

What the case study sacrifices in generalizability, it gains in empirical and theoretical richness. By focusing on one organization, I was able to collect a variety of data on the same phenomena at multiple levels of the organization and to do so over time. I performed non-participant observation of classrooms, department meetings, leadership meetings, and school-wide activities. I interviewed key informants about each change to gain perceptual and historical perspective. I also scraped the school's server to capture the documents being created during the school year as well as those accumulated in prior years.

This variety of data allows me to validate insights from one set of data by finding corresponding evidence in other types of data (Jick 1979). I am also able to trace processes across levels and units of analysis from individuals in meetings to the school as a whole and account for departments, committees, classrooms, and individuals. In doing so, I can and specify multilevel processes and cross-level interactions by teasing out micro and macro phenomena (Rousseau 1985; 2011).

Understanding how schools create effectiveness means studying purposive change. Thus, I examined how they defined success and what strategies they used to formulate, pursue, and achieve (or fail to achieve) that success. An integral part of the process of creating organizational behavior is planning, strategy, feedback, and other reflective and prospective processes by which goals are set, actions planned, and results observed. The trade-off is that this study cannot speak

to how schools become effective in other domains like civic education or reducing dropout rates. This study also cannot speak to how effectiveness emerges as a byproduct of other forces like competition or conflict.

Data Collection

I gained entrance to Achievement College Prep through an acquaintance who was a former teacher. Beginning in February 2014, I spent 20 months performing non-participant observation of routine activities at the school. I spent the first six months developing rapport with faculty and staff and establishing a strategy for what to observe. The academic year from August 2014 to June 2015 is what I refer to as my *focal year*. During this time, I sought to observe the life of the school from top to bottom over the course of a single school year. The subsequent four months from July to November 2015, I spent performing observation of events and meetings I believed went under-observed during the focal year.

I largely observed routine meetings of senior administrators, departments, classrooms, and whole-school activities. The purpose of the broad-spectrum observation was to capture formal organizational processes at multiple levels of the school. I limited my focus to three departments and the classes they taught: math, English, and history. I chose math and English to study the process of implementing Common Core and history acted as a foil where Common Core testing and standards did not apply.

I supplemented these observations by performing key informant interviews with senior administrators, supervisors, and teachers in the department I focused on. I asked interviewees how they came to Achievement, what jobs they have done at Achievement, and their experience with and opinions of the change initiatives on which I focused. These interviews enabled me to

locate facts, opinions, and experiences in the lived history and autobiographies of members of the school.

Finally, I supplemented the qualitative data with archival data from the school's servers. I scraped the school's share drive and archives from 2010 to 2015. Starting in 2010 gave me systematic data for all grades each year and provided data from before the studied initiatives had begun. For documents created by Microsoft office programs, I also saved document metadata including who created the document and who last edited the document and when. During the focal year, I performed weekly scrapes of the folders teachers and staff worked in. During these weekly scrapes, I saved copies of all documents which had been added or updated since the prior scrape and recorded any new metadata. This data helped triangulate the dynamics and facts I was uncovering in the other methods.

Analysis

I approached data analysis using an inductive case method (Eisenhardt and Graebner 2007). I chunked the school's changes into three initiatives: changing culture, building a professional community, and aligning to Common Core. For each, I answered a set of questions. What change are they pursuing? Why are they making the change? What was the strategy for achieving the change? How did the strategy play out during the focal year? What were the results at the end? These questions did not emerge from the observations themselves as would be case in grounded theory (B. G. Glaser and Strauss 2017); but were determined a priori from the research question.

The question of how the school pursued change requires a narrative answer. I constructed this narrative by reading my field notes and interviews and setting aside chunks that spoke to

specific questions. I translated these into a single, multi-vocal narrative of how the change occurred. These narratives form the core of this study. Finally, I also dove into the school's archives to answer specific questions or test certain hypotheses about these narratives. For example, I tracked the diffusion of Common Core standards in teachers' lesson plans from 2010 to 2015 to determine when teachers adopted these standards in their classes.

With these narratives in place, I returned to the literature to understand what these cases of change represented and attempted to reconcile existing explanations with the narratives here. The result are three cases of creating effectiveness within a single case study.

Results

What is Achievement College Prep doing that makes it more effective than other schools? It wasn't obvious when I walked in the door. It looked like any other school – cafeteria/auditorium with students in uniforms milling around. Classrooms along every wall with rows of desks facing the whiteboard and walls covered in inspirational quotes, school news, and other accouterments. Classes ran 20-30 students, four a day in the middle school eight in high school in 90- and 50-minute blocks respectively. There was nothing radical going on. They weren't giving every student a laptop or making them spin around in their chairs reciting Shakespeare. But when I asked them what they were doing to be more effective, they told me. And most any teacher or staff at the school said the same.

When I first met the school's director in 2013, she laid it out to me – building a good culture, hiring and training good teachers, and preparing for Common Core. These are bread and butter initiatives for any school. In chapter 2, I dig into why I think they're not radical and why, in general, charter schools have not been the engines of radical innovation that some of their

earliest proponents envisioned. But what I've learned is that you don't have to do something radical to improve student learning. You don't have to reinvent school to make it better for students. Many of the things ACP does can be done in other schools.

Culture is a best practice they're deeply committed to at Achievement College Prep. In the organizational and educational literature, organizational culture has been shown to be essential to performance. Among education thinkers, policy makers, and leaders; school culture is the cornerstone of leadership. It's no surprise that a school committed to high performance should be focused on creating a good culture.

In Chapter 3, I show how obsessed the school was with culture. For example, at the beginning of the 2015-2016 school year, they devoted seven days of teacher orientation to school culture. During my focal year, the cultural thrust was about becoming more "warm-strict." The school has full latitude to define and pursue a culture all its own. However, this degree of control and discretion came with a cost. The attempt to create a culture was a challenge, even when everyone seemed to understand and agree on what it should be. This form of change, which I call performative change, is in fact one of the most difficult kinds of change. It requires everything to go right in a way no other kind of change does. From my observations of the warm-strict change, it is unclear if their performance is due to their unified obsession with culture in general or their particular culture of high expectations and warm/strict discipline.

In Chapter 4, I dig into the what the school was doing to close the gap in their students' test scores. Faculty and staff at ACP were deeply committed to standardized testing as the measure of student learning. Their students had long-performed well on the state's standardized test but underperformed on college-entrance exams like the SAT and AP exams. They were convinced that the new Common Core standards were more rigorous and would be a better

measure of how prepared their students were for college. By the time I began my observations, faculty and staff were in an uncommon position of being both practically and ideologically committed to the Common Core, what I call "deep coupling."

Despite their deep coupling, they were not successful. My focal year was the first year the state allowed schools to take the new Common Core-based standardized test. Achievement was among the first to volunteer. Students fared poorly on the test, casting doubt on whether the school is truly effective or just teaching to the test. For teachers and staff however, this only confirmed that the new test was more rigorous and that, after seeing students pass state tests but fail college-entrance exams, they were finally getting more accurate feedback.

Finally, in Chapter 5 I look at the new things Achievement is doing for teachers. They have been mostly hiring teachers who were specially trained for no-excuses charter schools like them. All teachers were intensively coached by their department chairs and regularly observed by other teachers and administrators. And, every teacher worked on a one-year contract, thereby facing the possibility they could be let go every year.

This combination of intense training, coaching, and job insecurity led most teachers to leave within four years of joining the school. Those I spoke to reported working 60 to 80-hour weeks and, despite the work, always falling behind. And, given the cut in pay teachers take compared to local public schools and the lack of advancement opportunities, it makes sense that teachers would leave. This rapid hire-train-quit process makes the school an exemplar of what charter school critics have called "churn and burn." Given what we know about teacher effectiveness and seniority (i.e. that teachers do not reach peak performance until year five), it's surprising they do so well with such an inexperienced staff.

THE LONG ROAD AHEAD

Achievement is not doing anything radically new or weird. There are a lot of great things other schools can learn from and adapt to improve how they educate students. But they are not perfect or unequivocally effective. There's no miracle or foolproof best practice. From my data, I can see how the advantages of attending a no-excuses school on student learning are positive but weak in aggregate. To return to the metaphor I used earlier, they're only slightly better cooks.

So how is it that the momentous push to create excellent schools has accomplished so little? For all the energy we've spent creating excellent schools, the best schools we've managed to create are only slightly better than their traditional counterparts. Why didn't charter schools revolutionize learning? Why hasn't testing led to generalized increases in student learning? Why hasn't the prevalence of great research on culture and training led to breakthroughs in school and teacher performance? These questions point us to a more fundamental problem for education.

We are terrible at designing, testing, and scaling organizational interventions. In each chapter, I document how good research, well-intentioned leaders, and informed policy-makers founder on the rocks of organizational reality. Each change program I focus on confronts very different organizational challenges which make success practically impossible. The problem of intervention design is not simple or easy in organizations (nor is it simple and easy for any other science). In the final chapter, I discuss the problems of organizational design I identified at Achievement and how, if we are to make schools more effective over the next thirty years, we need to fundamentally re-think our approach to policy, research, and leadership.

CHAPTER 2

THE HOPE IN NO-EXCUSES CHARTER SCHOOLS

Consider the airplane. Modern airplanes fly at several hundred miles per hour from airport to airport carrying passengers looking to travel long distances in a short amount of time. The airplane made it possible for people to more efficiently conduct transnational business, transport goods globally, visit exotic locales, and relocate to another part of the world. It is a radical innovation that revolutionized human transportation. Yet, we do not send our children to school in airplanes or fly to the grocery store for milk. Flying is uneconomical for short trips and, logistically, we do not have the infrastructure for individuals to fly around their neighborhoods in personal planes. Despite its unambiguous status as revolutionary, the airplane has not monopolized transportation. Charter schools were supposed to be like airplanes.

There is a romance to the notion of radical social change – the idea that if we can start from scratch without rules, only big ideas, we can create something wholly new and wonderful. "Disruption," "revolution," and "paradigm shift" are all concepts we use to envision the ways in which society can become much better by making room for the very different. Chartering new, independent schools was presented as just such an opportunity. However, charter schools were not supposed to replace all schools. Their revolution was to create new ways to teach students underserved by contemporary schooling.

This is not what happened. Advocates of educational excellence glommed onto the idea quickly. Five years after the concept was invented, ten states and the federal government had passed laws authorizing the creation of charter schools. But the movement failed to bring radical social change to schooling. The few particularly innovative schools appear to have fizzled early.

Most successful charter schools in operation today look and perform like traditional schools. In fact, the majority of charter schools today are interchangeable with traditional district schools. Many more are little better than poorly run small business mired in mismanagement, poor leadership, and occasionally outright fraud. A few charter schools, colloquially called no-excuses charter schools, do seem to have found something that makes them more effective than their traditional school peers. Yet, these schools are not radically different in size, structure, or routine. They're not disruptors or revolutionaries. Instead of being the airplane of education, these charter schools represent something akin to a more fuel-efficient automobile.

Why did we reinvent the wheel (or automobile as the case may be) rather than lift off into a new age of flight in education? The fault lies not with a lack of innovation. Nor does it lie in a lack of competition. In the first half of this chapter, I argue the fault lies with our educational imagination. We're stuck with a particular idea of what schools are supposed to look like. Until we challenge these ideas, no amount of policymaking can revolutionize schooling.

What the success of no-excuses charter schools like Achievement College Prep tells us however is that there are still many things to improve upon within the existing model. On paper and in the classroom, Achievement looks like any other school and yet its students, most of whom are from disadvantaged backgrounds, perform well above their peers on state tests. In the second half of this chapter, I review the research on the impact of no-excuses charter schools and what the people at Achievement claim to be doing differently to understand where these schools claim to be making improvements before attending to them in the subsequent chapters.

THE CHARTER SCHOOL REVOLUTION

The Charter School Promise

The charter school movement was born out of the desire to find something that worked for underserved students. Its villain was the homogenizing process many saw in the standards and accountability movement through which schools across the country were adopting the same college-readiness curricula and testing. In his 1988 address to the National Press Club, Albert Shanker, then president of the American Federation of Teachers, America's largest teachers union, argued that state regulations, standards, and testing were creating a one-size-fits-all approach to education which benefited those "who are able to remember after they listen to someone else talk for five hours, who are able to pick up a book and learn from it" but "bypass[ed] the 80 percent of students [who can't]" (1988, 6).

Shanker's original proposal was for chartering pilot programs within schools. In his mind, a group of teachers with a great idea "could set up a school within that school which ultimately, if the procedure works and it's accepted, would be a totally autonomous school within that district" (12). He argued, "schools all across the country now, unfortunately, look very much alike. These schools will look very different, and they should follow certain guidelines that don't tell you what the school is going to look like" (12-13). In other words, rather than approving a particular type of school or strategy, officials would approve charters based on criteria for excellence.

Eight months after Shanker's speech, the notion of "charter schools" we know today was outlined by the Citizens League, a Minnesota-based think tank (Rollwagen and McLellan 1988). The Citizens League had been pushing for co-determined schools, those led by teachers and administrators and for improving schools in disadvantaged areas of the Twin Cities. Rather than

initially funding pilot programs within schools as Shanker suggested, they were the ones to recommend the state authorize the formation of whole schools. These "chartered schools" would be run by teachers, approved by the district, and certified by the state's department of education.

The charter school idea germinated in the span of six years from 1988 to 1994. The Citizens' League proposal for charter schools became the foundation for Minnesota's charter school law in 1991, the first such law in the United States. California passed its own version the next year. In the following ten years, 38 more states would add new charter authorization laws. Thirteen more states would pass authorization laws in the next fifteen years, bringing the total to 43 states with charter schools in 2015. Between 2000 and 2014, the number of charter schools increased 350% to 5,700. And the number of students enrolled in charter schools grew from 300,000 to 2.1 million, a seven-fold increase (National Center for Education Statistics 2014).

At the heart of the charter school promise were two sets of mechanism for creating and validating excellent education. First, charter schools would be given unusual regulatory freedom in order to encourage innovation (Nathan 1999). In charter authorization bills, states gave charter schools special waivers from many of the normal rules and regulations in order to enable transformative change. Second, charters would be held accountable for student learning. They could lose certification and funding if they failed to improve student achievement. Thus, those who improved learning would survive and those that didn't would close. This was the policy model, though the degree to which individual states and districts hewed to this model varied (Finnigan 2007).

For policymakers, charter schools would create a more effective public school system through the marriage of innovation and accountability. Charter schools offered school choice and competition on the one hand and radical innovation on the other (Berends 2015). Thus, the best

school should win out. This combination also produced rare bipartisan support for charter schools. Liberals supported them for their commitment to delivering effective education through public schools. Conservatives liked charter schools as avenues for school choice and a new source of competition for improving education.

The Charter Reality

Charter school policy should have led to a revolution in schooling in America. New groups were given license to organize schools however they wanted and teach in whatever way they wanted. The only limit was that they improve students' learning.[1] To organizational scholars, charter schools represent the ideal case for innovation: wholly new organizations created to be innovative and given waivers from laws and policies that typically regulate public schools. New organizations were to be created from scratch, eliminating the constraints of having to face the inertia of changing existing schools (Hannan and Freeman 1977; 1984). Moreover, the founders of these schools were to be a new breed of school leaders – institutional entrepreneurs with new visions for how schools could operate (Battilana, Leca, and Boxenbaum 2009). There would be a new, untapped market of public-school students who could choose the school they wanted and whose attendance was paid for by the state (Rao 1998).

The revolution never materialized (Berends 2015; Lubienski 2003; Preston et al. 2012). And, charter school policy appears to have been a wash academically. On average, students

[1] This claim is over-simplified somewhat. Different states passed charter school legislation with varying amounts of liberality (Finnigan 2007; Renzulli and Roscigno 2005). The Center for Education Reform provides an annual report card on the relatively freedoms given to charter schools by states. They measure regulation not only in terms of the specific freedoms charters are given from regulations, but also autonomy in authorizing charters, equal funding for charter and public schools, and rules for expanding the number of charter schools. The state in which Achievement operated was rated in the middle of the pack in the report card in 2014.

perform just as well on state tests in traditional schools as they do in charter schools (Berends 2015)[2]. However, charter schools vary significantly in their impact on student achievement. Many charter schools are poorly run or outright scams (U.S. Department of Education Office of Inspector General 2016). The vast majority are average. Some seem to perform exceptionally well. But before I get to them, let's answer the question we're here to answer. Why didn't the wave of charter school policies lead to the revolutionary improvement in education they were expected to?

REINVENTING THE WHEEL

Stability in Institutionalized Fields

Despite thirty years of experimentation, there has been no revolution in schools brought on by charter school innovation. Charter schools today look like every other school. But, charters were not always so mundane. In their study of Arizona charter schools, King, Clemens, and Fry (2011) found a variety of alternative models in the earliest days of charter school growth. For example, one school used the National Football League as a cultural template and another taught students reading, writing, and arithmetic through farming. These more creative schools seem to have been only marginal parts of the charter school population however. In their analysis of Oakland charter schools, Jha and Beckman (2017) found 20% of charter applicants included a such a "special theme" in their original charter application. However, the number of innovative schools is even smaller than that as Jha and Beckman included schools focused on arts and music as "themes" which are not quite as innovative as a farming-based school.

[2] There is disagreement in the literature over the effectiveness of charter schools. However, while there is substantial variation in the studies looking at this question, none show a large positive or negative effect on the whole. The debate is over the existence and size of a small effect.

Charter schools were generic because they and we, share a generic notion of what a school is. Reflecting on a decade of research on thousands of schools, John Goodlad (1984) tried to distill what he had learned about what makes schools effective. He found there was nothing that distinguished good and bad schools. They all generally looked the same. Good schools had good school culture, but so did many poor-performing schools. There were fewer effective teachers in bad schools, but there were still many effective teachers. He concluded that one would be hard pressed to distinguish the good school from the bad by just walking in the door or looking at numbers. Goodlad argued that the reason good and bad schools are so difficult to distinguish clearly is that they operate on a shared notion of *schooling*.

The notion of schooling has not changed since Goodlad's book. If you step into Achievement College Prep or any other charter school today, you'll find they look like every other school you've ever been in (Preston et al. 2012). Students are broken up into age-segregated classrooms of 20-30 students covering one subject at a time. They are evaluated through classwork, quizzes, and exams and, at the end of the year, if they have performed sufficiently well, they proceed with their classmates to the next grade. Teachers teach classes alone and occasionally in pairs within departments and grade levels. Principals and vice principals manage discipline, academic policy, school programs, and relationships with parents, regulators, and other external constituencies. There is orientation at the beginning of the year, graduation at the end, and predictable cycles of holidays, curricular units, and special events in between.

Schooling is what sociologists call an institution: a taken-for-granted set of practices, norms, and schema we use to organize and make sense of our world (Jepperson 1991). "Schooling" is the set of ideas about what school is from which individual schools take their

shape. Charter schools were no different. The institution of schooling limited the kinds of schools we imagined possible, even when we were trying to create policies to push those very limits. Achievement, like most other charter schools, did not challenge the central tenets of schooling.

Beyond the institution of schooling, education itself is an "institutionalized field;" a well-established area of social activity in which a stable set of regulatory, economic, and political forces engage in routine action(Stuit 2010). Not only do we all generally share a set of expectations about what a school looks like and how it works, but there are arrays of actors ready to enforce those expectations. King et al (2011) and Jha & Beckman (2017) both argue that new charter schools were not designed in a vacuum but were influenced by the education system already in place. In appealing to charter authorizers, accreditors, funders, students, parents, and teachers; charters had to be distinctive enough to warrant taking a gamble on but could not be so radical as to be unrecognizable as schools. In other words, the institutionalized field forced charter schools into a Goldilocks problem of being just innovative enough.

The type of innovation charter school operators offered was different from disrupting schooling. They did not propose a new vision of education but one that fixed what was wrong with traditional schools. Where public-school classrooms were unruly, theirs would be well-mannered. Where public-school teachers were ineffective, they would hire the best and brightest and relentlessly push them to get better or get out. They would improve test scores by teaching longer, faster, and harder.

Even if most proposed charter schools were not radically innovative, why didn't the few with bold, effective ideas rise to the top? One reason is that regulations for who can start a charter school varied widely leading many schools to be founded by people unprepared to run a

school (Stuit 2010). Freedom cuts both ways. Another reason is what's called the "liability of newness" (Stinchcombe 1965). For all the advantages new organizations have in being different, they are also faced with substantial challenges. New and innovative organizations require extra work to define new roles, build new relationships, and acquire resources. School founders with a great idea still must learn to make payroll, recruit students, and hire teachers.

Finally, the mechanisms of accountability and competition meant to drive innovation upwards have yet to materialize. Of all charter schools ever opened in the U.S., 85% are still in operation today. Of the 15% that closed, the large majority (66%) closed for financial and managerial reasons. Only 19% closed for academic reasons (The Center for Education Reform 2011). In other words, while many charter schools perform worse than schools in their district and the majority are interchangeable with district schools, only 3% of charter schools have ever been closed for poor academic performance. Given districts' unwillingness to enforce accountability, it should be no surprise that the average charter school is just that, average. Moreover, this lack of accountability means charter schools appear to pose no threat to traditional schools. Districts and school leaders report little sense of competition with charters and the presence of charters does not seem to affect traditional schools (Arsen and Ni 2012; T. M. Davis 2013; Zimmer et al. 2009).

The Exception

Despite the aggregate news about charter schools, there is substantial variation. The no-excuses charter school is the one type of charter school that appears to be systematically more effective than their public school counterpart (Betts and Tang 2019; Gleason et al. 2010; Center for Research on Education Outcomes (CREDO) 2009; Hoxby, Murarka, and Kang 2009;

National Center for Education Statistics 2014). These schools target disadvantaged students and emphasize high academic expectations and strict behavioral control. They argue that given the right ethos and high standards, disadvantaged students can achieve just as much if not more than their well-to-do peers.

A few studies have investigated what features make no-excuses schools so much more effective than others. Some of the studies that look at the individual policies and characteristics of schools find that longer school days and intensive tutoring accounts for increases in student learning (Gleason et al. 2010; Center for Research on Education Outcomes (CREDO) 2009; Hoxby, Murarka, and Kang 2009; National Center for Education Statistics 2014). In other words, they provide *more* school.

However, no-excuses schools have been shown to do three other things leading to improved test scores – focusing on achievement, student discipline, and teacher development. The charter schools that perform better do three things. First their teachers and staff say they set high expectations, expect students to try hard and complete all assignments, and expect students do well on state tests. Second, they impose strict standards on students' behavior. Third, they provide coaching and feedback to teachers (Berends et al. 2010; Dobbie and Fryer 2013; Furgeson et al. 2012). These are the three factors I focus on in my study of Achievement College Prep.

ACHIEVEMENT COLLEGE PREP

Achievement College Prep occupied a small, brick box surrounded on three sides by the staff parking lot, which doubled as a playground. Its periphery was ringed by a black, wrought iron fence, itself ringed by the city's sidewalk, kept busy by a patch of mom and pop restaurants,

corner stores, retail banks, repair shops, and a library nearby. On the small front lawn, "Achievement College Preparatory School" was printed on a banner spread taught between two wooden posts planted in the ground. The name of the building's original tenant was engraved in cement high above the front entrance: "St. Mary's Catholic School."

Before the charter school movement began, Catholic schools were believed to offer the key to building good schools. Right before the publication of *A Nation at Risk*, James Coleman, the famous school researcher and Chicago sociologist, had initiated what was at the time the second great debate about school effectiveness with the publication of the *High School and Beyond* study (James Samuel Coleman, Hoffer, and Kilgore 1982). The study found that students in Catholic schools, especially those from socio-economic disadvantage, learned more than students in public schools.

This was only the most recent incarnation of a twenty-year debate among academics over whether schools really had any effect on education or social inequality. This was the earlier debate Coleman kicked off in publishing his *Equality of Educational Opportunity* report (1966). In what was simply called the Coleman Report, Coleman and his colleagues asserted that the primary predictor of how a student did in school was how their parents did in school. Schools played little role in social mobility. The ensuing years saw an explosion of research into whether any schools were effective and what, if anything, makes one school more effective than another.

Catholic schools did see an increase in enrollments through the 1990s and early 2000s. Yet, the push for vouchers found limited political success. Shifting demographics, changes in the teaching workforce, and church economics caused the number of Catholic schools to collapse (P. Meyer 2007). Today, they currently stand at half the enrollment they had in 1950. Charter schools have taken their place in the imagination of education policymakers, researchers, the

media, teachers, parents, and students. Some, like Achievement College Prep, have even taken their place in their old buildings like organizational hermit crabs.

The front door to Achievement remained locked throughout the day except for the thirty or so minutes in the morning when students would arrive by public transit, school bus, car, or on foot. If you arrived with the students in the morning, you would unfailingly be met by a teacher assigned to greet all comers in the foyer and the din of students milling around before class. If you didn't come in when students arrived, you had to ring the buzzer and be let in.

Inside, the school had three floors including the garden level housing the middle school and the first and second levels housing the high school. Each floor was a large rectangle the size of a football field with classrooms along each wall and open space in the middle. The main entrance to the school opened into a short hallway with an office on each side that emptied into the open area on the first floor which was occupied by tables and surrounded by bulletin boards. This was used as the high school cafeteria and auditorium with a stage along the opposite wall doubling as the school's library.

The classes are generally split up by subject area. On the first floor, the science classrooms were arrayed across from the English classrooms like the broadsides of opposing ships. On the second floor above English were the History classrooms rivaled by Math across the gangway. Administrative offices, the reception office, a conference room, the dean's office, and a faculty lounge occupied the no-man's land along the front wall.

On the garden level, the middle school rooms were arranged by grade with eighth on one side, seventh on the other, and sixth at the far end below the stage. In the middle of it all was the middle school's cafeteria where students congregated before school and ate breakfast and lunch

according to their homeroom. Rounding out the garden level were a faculty lounge, the nurse's office, and small rooms for tutoring and meetings at the ends of the floor.

Achievement College Prep is one of the best charter schools in a region known for effective public schools. But one wouldn't know it by walking around. A brief tour of the school and its ordinariness was clear. So too was its social geography.

The Social Organization of the School

The school employed roughly sixty full-time and part-time staff per year and taught approximately four hundred students. Students enrolled via a lottery at the start of middle school or high school. In middle school, students took courses in history, math, reading, English, and science. In high school, they substituted Latin for reading and added an ethics sequence. Class periods also shrank from ninety minutes in middle school to fifty minutes. Special periods for art, gym, and reading were interspersed throughout the week.

The classroom structure was ordinary as well: a teacher led a group of 20-30 students through a lesson each day, sometimes with an assistant. All of the students were in the same grade and the lesson fell within the narrowly-defined disciplines of English, math, science, or history. Students advanced as a cohort from grade to grade based on their performance on teachers' tests and occasional standardized tests. As they advanced, they took classes in the same four fields over again, English, Math, Science, History; with different foci, skill level, and teacher.

Teachers' work was also ordinary. Middle school teachers taught two 90-minute classes per day while high school teachers taught four 50-minute courses per day. Teachers met with their department weekly and with their grade level biweekly. Teachers also met individually with

their department chair once a week for an hour to discuss their professional development and upcoming lesson plans. All staff had a homeroom period at the end of the day and participated in weekly, hour-long community meetings with faculty and students in the middle and/or high school and the whole school. Finally, teachers had two or three duties each week such as lunch monitor, reading proctor, detention, or student arrival and departure. The whole staff met once a week for the faculty meeting where all staff share news, discuss school-wide issues, met as high school and middle school teams, or did professional development activities.

Roughly 70% of its students were African American and 30% were Latino. Ninety percent come from low-income families. The staff on the other hand were predominantly white, female, and right out of college.

What's Really Going on?

You can be excused for failing to see what made Achievement College Prep different from any other school in the country. On its face, there was little to suggest that it was one of the most effective schools in the state. Yet, students at Achievement performed much better than their peers. During their first years at the school, middle school students' test scores were middling compared to the state. Yet, by high school their state test scores exceeded some of the most elite schools and were comparable to the area's high performing exam schools. Every student who graduated from Achievement had gone on to college. The year I began my observations, the school won an award from a local foundation as the best public school in the region.

It looked like any other school anyone has ever been in. Yet Achievement is able to do something that has eluded policy-makers, researchers, school leaders, and educators for forty

years – make good on the American promise of an excellent education regardless of the conditions into which one is born. How is this possible?

Before I began this study, I had to meet with administrators to negotiate access to the school. For an observational study, getting access to an organization primarily means getting the consent of the school's director, Kim. Kim began her career teaching math in well-to-do middle schools but transitioned to not-for-profit management, attending an elite MBA program and then working for a nationally-recognized management consulting firm before being brought on to succeed the school's founder as its second executive director. I met Kim in the humble meeting room at the entrance to the school typically reserved for parents and small committees. It was a room with ten chairs and a table for six with one of the few southerly facing windows allowing in direct sunlight.

I explained to Kim that I was interested in studying how they were trying to change education. She was excited to tell me what they were working on: improving the school culture, transitioning to Common Core, and trying to build a professional community. At the time, I got excited that there was so much change to be studied at the school. In a way, I felt like Kim handed me my dissertation right there. All I had to do was come in and watch them work their magic.

That day I didn't make it much further than that front room. But as is true about much of Achievement, you don't have to get too deep to find out what makes it tick. On paper or just walking around, they look like every other school. But once you start talking to them, the list of programs and initiatives they're pursuing to make themselves effective are front and center.

THE NO-EXCUSES SCHOOLS MOVEMENT

To say that charter schools reinvented the wheel means that charter schools started as relatively novel or independent schools and adopted mainstream features of schools over time. Where Jha & Beckman and King et al examine this process at the field level, this study looks at how this process of convergence occurred at Achievement College Prep. The no-excuses movement was and is based on the idea that new schools need to be built from the ground up to escape the trappings of traditional schools. At their birth, Achievement College Prep and other no-excuses schools differentiated themselves from traditional public schools by pointing to a new model for educating disadvantaged students. This model highlighted tight behavioral standards, high academic expectations for students, performance-based contracts for teachers, and an extended school day.

The no-excuses movement started early in the charter school movement with its first schools established in the mid to late 1990s: KIPP in 1994; Uncommon Schools in 1997, YES Prep in 1998, Achievement First in 1999, and Aspire Charter Schools in 1999 (Cheng et al. 2017). These schools were founded independently of one another but shared the premise that disadvantaged students needed high expectations to achieve. What was wrong with traditional public schools serving economically, racially, and culturally disadvantaged students was that these schools approached their students as if they were uneducable (Carter 2000; Thernstrom and Thernstrom 2003). Thus, at the heart of the no-excuses movement was the belief that educational inequality could be solved by good school management (Trujillo, Scott, and Rivera 2017).

The no-excuses movement was not christened as a field until perhaps Thernstrom & Thernstrom's book with the name and outlining many of the same values expressed by these founders and pointing to them as paragons of the model. Few schools considered members of the

no-excuses movement endorse the name formally and rarely use it in their promotional material. There is no formal organization coordinating the movement or setting its agenda. KIPP, Uncommon, and Yes Prep run their own independent charter management organizations while individual no-excuses schools participate in regional networks of similarly-inspired schools. The movement is dispersed.

Beyond these shared beliefs, no-excuses schools tend to share a set of practices (Furgeson et al. 2012; Golann 2015; Thernstrom and Thernstrom 2003). One is a longer school day and longer school year. As Thernstrom & Thernstrom (p46) put it, "the children at these schools are playing catch-up with their more advantaged peers. They need and get more instructional time than regular public schools." Another distinct practice is broken windows discipline characterized by harsh punishments for mild rule violations. Finally, high expectations mean a focus on increasing student test scores and expecting all students to graduate and attend college. This list is not exhaustive but exemplary of the policies these schools put in place to distinguish themselves from traditional schools.

The aspirations of the no-excuses movement were grandiose – provide disadvantaged students with the skills, knowledge, and experiences equal to their advantaged peers. Yet, for the most part, the techniques, policies, and structures shared by movement members are neither atypical nor unprecedented. They did not abolish grade levels or disrupt the curriculum. What they see as innovative and effective is a relatively narrow, but highly specific set of inventions like a novel school culture, an alternative approach to teacher training, and a focus on test scores and college attendance.

WHAT IS NEW AT ACHIEVEMENT?

Achievement College Prep was founded in the no-excuses mold. But the school is continuously working to make itself better. Near the end of my fieldwork, I listed out the many initiatives I observed in order to ensure I was covering the most important ones in the most robust way. One result of this was lumping a variety of changes into the three initiatives. Another was recognizing the variety of change initiatives I could not cover in this study.

The initiatives I chose to focus on were those that were the most important and substantial initiatives to the school. Those I did not include had one or more of the following characteristics. First, there were many changes I didn't include because, while important, they were relatively narrow in scope. For example, the school changed its homework policy to encourage teachers to use formative assessments and grade for completion in addition to grading for accuracy.

Second, there were changes that were part of the initiatives I study which I do not focus on because they were either under-developed during my observations or they were too ancillary to the main strategies being pursued. For example, one change to student culture involved creating extra-curricular activities and student government as ways to build school spirit and student engagement. This was an important shift to members of the school, but it was not seen as central to the culture of the school when compared to the changes they were making to their disciplinary policy.

I also did not study all of the new things that made the school effective. For example, the school sends all of its students to college. That's a phenomenal achievement which they accomplish through the concerted effort of teachers, guidance counselors, parents, and students working to ensure that all students create college applications ahead of time, submit many

applications, apply for financial aid, and ultimately enroll in a college or university. However, I was not present when they created the system and cannot say what it was that allowed them to become effective at getting students into college. I wasn't able to observe the creation or change process.

No-Excuses Culture

Culture in schools has had two widely-understood meanings. The first is enculturation. Schools enculturate students into some beliefs, values, and norms about the world. This process, often referred to as the "hidden curriculum," is often observed to be subliminal, indirect, or a byproduct of other school functions. It is unintentional. Perhaps the most famous analysis of the school enculturation is Paul Willis' book *Learning to Labor* (1977) in which he observed how working-class students' experience of schooling taught them to resist education, accept authority, and prize manual over mental labor.

The second is the culture of the school itself: the beliefs, values, and norms holding its members together, commonly called "school culture." School culture is a type of organizational culture – "the way we do things around here." Historically, it has been treated as intransigent (Sarason 1971). More recently, school culture has been seen as malleable by leadership. It is a tool for school leaders to shape coordination and control among staff and students (Deal and Peterson 1999).

No-excuses schools were founded on a model of education that made school culture and student enculturation work together. They made the hidden curriculum explicit. Schools should instill in disadvantaged students a "toolkit" of skills and habits to attain the academic and personal successes of their more advantaged peers (Thernstrom and Thernstrom 2003). To

bestow this toolkit, school staff must adopt and abide by a set of values, norms, and beliefs that establish high expectations for students – a culture of no excuses (Lake et al. 2010). One problem with this vision has been how this has shaped these schools' disciplinary policies.

No-excuses schools set high expectations that instill habits of success in part through strictly-enforced behavioral codes. The no-excuses philosophy asserts that schools must create and strictly enforce rules like sitting up straight, bringing materials to class, looking at whomever is speaking, and keeping one's school uniform tucked in. (These rules are in addition to common rules against being disruptive, fighting, cheating, and possessing weapons.) Strict enforcement of these rules leads to high rates of punishment and punishment is multiplicative. At Achievement, not having a pencil, slouching in one's seat, and dropping a piece of paper each earn a student a demerit. Three demerits taken together automatically get a student sent out of class. These minor offenses also quickly multiply into suspensions.

In no-excuses schools like Achievement, annual suspension rates can be as high as 50%. (The national average is roughly 5%.) These sensationally high rates have stirred national controversies over whether charter schools compound students' disadvantages by further excluding them from education, marring their academic record, and inflating the school's academic numbers by pushing poor performing students out. Joanne Golann (2015) has challenged the notion that such discipline teaches self-control by arguing that prioritizing control leads students to not challenge authority, withhold novel ideas, and keep to themselves rather than taking the initiative.

As I discuss in Chapter 3, Achievement College Prep responded to this controversy by rethinking their approach to culture. They gave up one of the main characteristics that made them distinct: harsh punishments for minor rule violations. They also gave up on some of their rules

and gave teachers more discretion in disciplining students. As a result, Achievement cut the rate of suspensions to a fraction of what they were.

However, they remained committed to their core belief that strict discipline was necessary. During my focal year, they were moving away from a culture of discipline that was harsh and strict to one that was warm and strict or just "warm-strict." They would maintain high expectations by strictly enforcing rules but in a way that treated students with respect. They would not punish students who accidently dropped a pencil or left their shirt untucked after coming from the bathroom. Instead, they would "read students intent," punishing those who knowingly or carelessly broke the rules. They believed this new approach to culture would build trust among students and teachers and eliminate unnecessary discipline like sendouts and detentions.

Common Core

While charter schools have been able to garner exemptions from many school regulations, one they largely could not be exempted from and which many did not want exemption from was standardized testing. Many charter schools were founded to produce students who could excel at standardized tests with the belief that tests scores reflected real student learning. No-excuses schools have been especially driven by these tests. Their express goal is to close the gap between rich and poor, black and white students and they measure this gap by the average difference in students' state test scores, graduation rates, and college attendance. At Achievement, staff measure their success by how many students pass the state test, graduate, get accepted into colleges, receive financial aid, attend a college, and graduate with a four-year degree within six years.

This system of metrics is based in the same school accountability regime formed after *A Nation at Risk* tying state standards, standardized tests, and state resources together. Common Core represented the next iteration on this regime – creating more rigorous standards for students to meet requiring a new test to measure student learning and evaluate school performance. The Common Core State Standards were developed by the National Governors Association in 2009 and adopted by many states that same year during President Obama's Race to the Top. The first standardized tests meant to measure student learning on Common Core standards were developed and tested between 2013 and 2014 and, during my focal year, the first states began using these tests to assess student learning and school effectiveness.

Whereas the shift in school culture and discipline were driven by criticisms from within and outside the no-excuses school movement, changes to school curricula and testing came from the government. However, teachers at Achievement did not merely align their curricula to the new standards. They internalized the worldview contained in them. They bought into the theory behind Common Core – that learning-how rather than learning-that represented a deeper, more generalizable, and more collegiate form of knowledge. Unlike schools before them, they coupled deeply to the Common Core.

Creating Effective Teachers

Charter schools were to improve public education in part by changing the way teachers were recruited, retained, and compensated. The ideal charter school was supposed to recruit the best and brightest to become teachers, provide intensive on-the-job training for these wunderkind, and reward the effective and fire the ineffective (Hoxby 2002; M. J. Podgursky and Springer 2007; Neal 2011; Furgeson et al. 2012). To pursue this ideal, charters were given

waivers to recruit teachers from non-traditional training programs and exemptions from collective bargaining agreements with teachers unions which regulated teacher compensation, retention, and work hours (M. Podgursky and Ballou 2001). Charter schools in general have pursued this ideal (see Furgeson et al (2012) for an overview) and no-excuses charter schools in particular have taken it to the extreme (Lake et al. 2010, 5, 58; S. F. Wilson 2009)

Given the blank check, charter schools like Achievement have been much more likely to hire people with unconventional backgrounds. Charter school teachers in general are much more likely to have alternative degrees and teaching certificates rather than the four-year Bachelor's degree in education. A near unanimity of teachers at Achievement studied something other than education in college and many had master's degrees, a few in education. Most teachers at Achievement were recruited through alternative teacher training programs like Teach for America and a homegrown training program I pseudonymously call the No-Excuses Teacher Training (NETT) program.

Intensive on-the-job training compensates for the lack of classroom experience that teachers comes through this non-traditional path have. Teachers in charter schools tend to receive more observations, spend more time in professional development, and receive more reviews than those in traditional public schools. This is part and parcel of the belief among charter school operators, including Achievement's leadership, that the best way to learn to teach is by highly mentored teaching. As I mentioned earlier, at Achievement, teachers were observed weekly by their department chairs and participated in up to five formal professional development activities each week.

Contracts for teachers at charter schools vary widely but typically offer much less salary than traditional public-school jobs. At Achievement, all teachers were hired on one-year

contracts which could be renewed each year based on the teacher's performance in the classroom. Despite the annual threat of losing their job, teachers' contracts were almost always renewed.

Unfortunately, no-excuses charter schools experience what's known as "churn and burn" where most teachers work long hours, burn out, and leave the school after only a few years. While I never conducted a survey of teachers to estimate their average work hours, those I interviewed reported 60-80 hour work weeks and this was understood as normal if regrettable. During my focal year, the average tenure of a current teacher was only two years.

At no point did anyone at any time say this system was the way they wanted to create an effective teaching staff. Instead, there was an implicit assumption that the ideal teachers at Achievement was a type-a over-achiever which guided the many disparate decisions about hiring, pay, hours, training, and retention. This assumption is not shared by staff in traditional public schools and is not a premise of the mainstream institution of schooling. Instead, it was imprinted in the organization at its founding causing churn and burn to emerge from an otherwise loose collection of practices

WHAT'S TO COME

Over the next three chapters, I outline what the school tried to accomplish with these innovations and identify whether and how they worked over the course of my focal year. I find that each one is deeply embedded in the strategy of the school and has been for years. These innovations are real and understood to be important by everyone at the school. They are also mature strategies that have been part of the school's daily life for years and, in their view, critical to the success they've already had and to their prospects. Yet, during my focal year, they

encountered severe challenges and often failed to achieve the goals they set. In fact, some of their failures suggest that the studies showing that no-excuses schools are effective may be wrong. However, it is well beyond the scope of this study to decide whether these strategies are in fact effective and if they are the reason why no-excuses schools are more effective. My goal is to understand how these purported innovations are actually playing out in the life of the school to help us understand whether they are *plausibly* related to school effectiveness. On this question, the answer is unfortunately mixed.

CHAPTER 3

WHEN CULTURE IS THE STRATEGY

INTRODUCTION

The ideal case of change is when science provides strong evidence that a change will work and practitioners go about implementing it. For example, researchers study lots of students to determine the most effective ways to teach math. Educators then implement these methods into their teaching and students learn more math over time. This is not just the ideal for education. One tenet of organizational scholarship and all of science is to create such practically-relevant knowledge (Douglas 2009; Nicolai and Seidl 2010).

I use the term performativity to describe the process of adopting a scientific prescript into an organization to effect change. The term comes from science and technology studies, which adapted it from philosophy. In this context, it means using a theory as the basis for one's practice (the intuition being that practitioners are "performing" the theory).

As a type of organizational change strategy, performativity entails both clarity and control. The theory provides clarity – a set of key concepts, mechanisms for how to use these concepts to affect action, and a set of expected outcomes resulting from these actions. But theoretical clarity is not enough. Performers must also create the conditions in which the theory can work. Understanding aerodynamics does not equip one to fly. You must have designs, factories, test facilities, and pilots. In this and other cases of performativity, organizations need to control the resources, people, and technology to effect the change.

Performativity represents a certain extreme case of maximal clarity and control. If we are to understand how we have failed to improve education, we must understand this type of case. The majority of research on performativity has been on cases where we know the outcome already. Here is a case where the outcome is known only in the short term. We do not yet know whether the no-excuses theory will ultimately hold up. This is important because practitioners and applied researchers rarely have clear knowledge of success or failure when they make decisions about whether to continue or abandon some theory.

Based on my research, there is little prospect for educational and organizational scholarship to form the basis for performative change. At each step in their attempt to create a warm-strict version of no-excuses culture, the theory became vague, murky, or inconsistent. Errors in implementation are incommensurate, making it impossible to identify the source of problems or their potential amelioration.

If we compare culture at Achievement to a medical intervention, whereas the process of implementing a culture was idiosyncratic and ultimately bootstrapped at Achievement, medical interventions have clear procedures like dosage schedules. If the intervention does not work, there are treatment progression strategies: increase the dosage, rotate versions of treatments, combine the treatment with supplementary interventions. At Achievement, errors in implementation compound to the point it becomes impossible to decide if the theory is wrong or just the particular implementation strategy.

The Origins of No-Excuses Culture

Achievement's leaders did not pick up an issue of the *Academy of Management Journal*, *American Education Review*, or the *Harvard Business Review* one day and say, "Hey, we need to

get a culture." They were not founded by scientists who had made a new discovery in teaching children through culture. Instead, their culture developed as part of a broader social movement led by educational entrepreneurs looking to redefine schools.

Culture had long been a subject of study in education and a central concern for scholars studying the effectiveness of schools. For example, Coleman and Johnstone (1963) looked at the culture of academics and athletics among adolescents. Bowles and Gintis (1976) and Willis (1977) looked at the ways in which the school produced inequality by reproducing class and work culture. In the 1980s and 1990s, organizational scholars, who had begun studying culture as a means for improving businesses' bottom lines, began looking at culture as a way to improve schools (Deal and Kennedy 1983; Deal and Peterson 1999; Teddlie and Stringfield 1993). This work positioned organizational culture as one of the main levers affecting school success. This quickly translated into a focus on studies in how to *create culture* through leadership (Campo 1993; Firestone and Wilson 1985; Leithwood and Jantzi 1990).

It was at this time that no-excuses charter schools were founded (Cheng et al. 2017; Renzulli and Roscigno 2005). Important too was the founding of Teach for America in 1989 which presented similar views of the educational landscape and whose alumni went on to found many of these schools, lead school districts, and write educational policy (Angrist, Pathak, and Walters 2013; Lake et al. 2010). These were the initial institutional entrepreneurs who seeded the movement.

These institutional entrepreneurs adapted theories of culture to understand and prescribe their own solutions for their specific problem. They believed that disadvantaged students performed poorly in school because schools accepted and enabled low expectations. They argued setting high academic and behavioral expectations would close the achievement gap. They built

no-excuses schools and put organizational culture at the center of their philosophy (Thernstrom and Thernstrom 2003).

In sum, the theory that organizational culture influenced school effectiveness came to prominence at the same time that a cohort of institutional entrepreneurs in education sought to address systemic social inequalities by creating more effective schools. These entrepreneurs, like generations of managerialist reformers before them (Cuban and Tyack 1995; Jal Mehta 2013), transposed the solutions offered by management into the problem space of public education and built schools on that basis. In doing so, they attempted to perform their theory to realize a culture of high expectations that would teach discipline to students and close the achievement gap.

PERFORMATIVITY THEORY IN EDUCATIONAL ORGANIZATIONS

As traditionally understood, performativity is a type of social construction – the application of a social construct causes that construct to become real or true (Gond et al. 2016). This is different from a self-fulfilling prophecy in that self-fulfilling prophecies become real because people believe them. Performative concepts become real because people create things that make them true (Callon 2007). Evidence of performativity has been wide-ranging from biology (Latour 1987) to economic theory (Callon 2007; MacKenzie 2003; MacKenzie and Millo 2003) to gender (Butler 1988). Only recently have organizational scholars begun looking at the extent to which organizational theories are performative. In education, the concept of performativity is wholly foreign.

Performativity in Organizational Science

The proposition that organizational theories are performative seems an obvious one. Corporate governance, organizational structure, strategy, and culture result from choices

organizations make and these choices are very often informed by the latest research publicized in outlets like the *Harvard Business Review*. In doing so, organizations can create the reality academics theorize (Ferraro, Pfeffer, and Sutton 2005; 2009).

Despite its goal to provide practical theory and the tight relationship between organizational science and organizations, there have been few assessments or case studies of how organizations attempt to perform theories in the field. For example, D'Adderio and Pollock (2014) show how theories of organizational modularity influenced organizational design in an electronics manufacturing firm. Other cases include advertising companies using auction theory to create auctions to sell advertisements online (Glaser, Fiss, and Kennedy 2016) and using an honor code to instill theory-laden values (J. Gehman, Trevino, and Garud 2013).

With collaborators, Dobbin has argued that shareholder value and agency theory were at the heart of corporate governance reform in the last decades of the twentieth century (Dobbin and Jung 2010; Dobbin and Zorn 2005). In the 1970s, organizational scholars studying principal-agent theory portrayed the central problem of organizational action to be how owners (the "principals") could shape the behavior of managers ("their agents") (Jensen and Meckling 1976; Fama 1980; Fama and Jensen 1983). Their solution was shareholder value theory which argued the central role of the firm was to make money for owners and recommended a set of policies and practices like performance-based pay for executives and debt financing to ensure this goal was met. Shareholder value subsequently became a prominent concept in the field, organizing the way many investors, analysts, and executives articulated their interests. Despite the dominance of the concept, the outsized influence of executives over their corporate boards led companies to only adopt the strategies that favored executives creating a perversely incentivize

governance system with high risks. While these measures performed in the short run, increasing stock prices, they exacerbated the financial crisis of 2008.

Performativity in Education

Research into the links between theory and practice are part and parcel of education. Education researchers perform interventions to see if those interventions can influence outcomes in schools. However, most study vague policy initiatives like charter schools, bussing, or standardized testing as proxies for theories. Few study specific programs with defined implementation and evaluation criteria. Fewer still use randomized control trials for interventions, the gold standard for proving a theory has interventive power.

However, the concept of performativity and its attendant insights have been absent. For schools to perform a theory like organizational culture, they cannot just claim to embrace it. Like the vaccines and needles in medicine, they must adopt its social and material technologies – the presentations and values, explicit beliefs and norms, schedules for enculturation and rules for violation. What is missing from the study of educational interventions as performances of theory is an understanding of the tools for program design, implementation, and diagnosis for performing theory.

Ironically, the process of designing, implementing, and diagnosing has often been papered over in existing work in performativity. As I mentioned, existing research has focused on cases in which the results are already known. The design, implementation, and diagnostic processes have already been completed. During my time in Achievement, I was able to see the uncertainty unfold across time. Good designs unraveled. Processes did not go as planned. The

diagnosis was murky. In other words, there was a performative process going on that has thus far been unexplored anywhere.

A MODEL OF PERFORMATIVE CYCLES

Michel Callon pointed to such a process of testing, tweaking and retesting, calling it "performation" (Callon 2007). Most researchers gloss over the performation process, presenting it as an arrow from theory to implementation to performance. Yet, during this process actors are answering the essential questions: Does the theory work? Is this a felicitous implementation? Is this a good performance? Why are bad performances happening? Should we abandon the theory?

In the remainder of this chapter, I open the black box of performation as it occurred at Achievement College Prep. In the next section, I briefly describe how the concept of organizational culture was translated from academia by individuals looking for prescriptions for organizational success. I then outline a five-stage performativity cycle based on my observations of Achievement. Each of these stages presents untheorized pathways for theories and the schools implementing them to succeed or fail.

Organizational Culture as a Performative Theory

Culture is one of many domains in which organizational theories have crossed over into practice. The cultural turn in organizational analysis in the 1980s translated into an explosion of academic research and popular management writing (see Weeks (2004) for an analysis and Chatman and O'Reilly (2016) for a recent review). Today, culture is a mainstay in management classes, the mainstream management press, and the board room (Graham et al. 2016). The cultural theories that develop at the interface between researchers and managers are almost always constructed from the perspective of organizational leadership with the concomitant

prescriptions for how to create culture "from the top down"(Schneider, Ehrhart, and Macey 2013). The explicit purpose of designing and creating an organizational culture is to improve organizational performance (Schein 2010).

This approach to culture situates it as an input into organizational life determining performance (Harrison and Carroll 2006; O'Reilly and Chatman 1996). In this view, culture is the set of beliefs, values, and norms endorsed by the organization and shared by members. This culture-as-an-input affects organizational outcomes by enabling members to coordinate their behavior to accomplish goals. This contrasts with a second way culture is traditionally approached in organizational analysis: as the emergent output of organizational life (Chatman and O'Reilly 2016). In this approach, organizational behavior drives culture. Culture is made up of the beliefs, values, and norms members express and exemplify *as a result of* their experiences working together in the organization (Martin 2002; Smircich 1983).

To approach culture as performative means treating it as both an input and an output. Organizational leaders design a set of beliefs, values, and norms for the organization and members express and exemplify them. It is engineered by leadership and emergent in the lives of its members. Few studies examine culture as both an input and output (see Kunda (2009) and Weeks (2004) for exceptions). This study suggests that as culture has become a common part of organizational leadership, a theory of culture must include top-down and bottom-up processes.

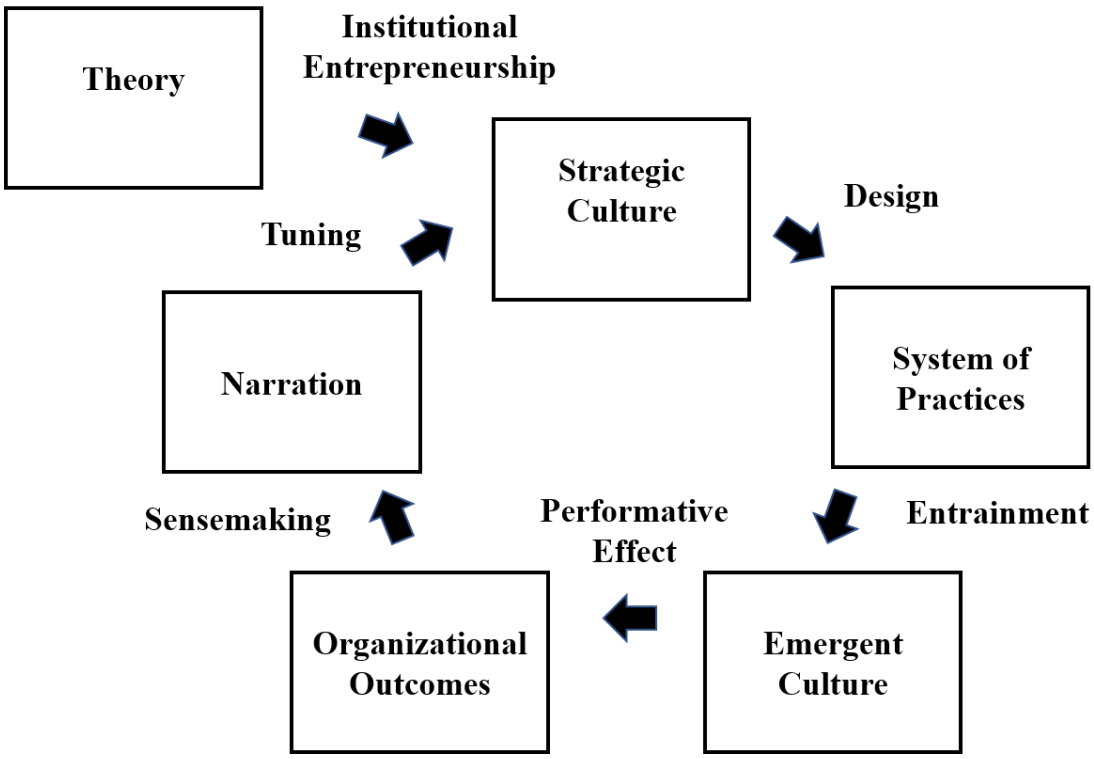

Figure 3.1: Model of the Performativity Cycle for Culture

A Model of the Performativity Cycle

Figure 3.1 contains the model that best captures the process by which Achievement attempted to create an organizational culture. I treat the performativity cycle as an iterative cycle of trial and error. It begins when a school adopts a model for organizational culture from institutional entrepreneurs who have done the work of translating a theory into a set of practical ideas. The school creates a strategic model, translating the generic recommendations into a set of actions they believe they can take to realize their own culture. They design organizational routines and procedures to carry out these actions, what I call the system of practices, constituting what the organization tried to do. In going through with these systems, a culture emerges resulting in some observable effect on the intended and unintended goals of the organization. Organizations interpret these outcomes by building causal narratives of what

happened and why. Based on these narratives, they update their model, adjust their design, and try again. While this model of the performativity cycle is articulated in terms of culture, it could be applied to other types of organizational performativity.

The success of performative change hinges on each stage in the cycle and each stage presents its own unique set of challenges. For example, the strategy must be realistic and coherent. Outcomes must be observable and relevant to the organization's targeted goals. And, because each stage feeds into the next, any failure in one part can create a failure downstream. If strategy is incoherent, the system of practices can be disjointed, redundant, or non-sensical.

Finally, contrary to what scholars of performativity have previously posited, success and failure are neither a continuum nor a binary. Instead, there are different categories of success and failure depending on which stages went well and which did not. I argue the culture counter-performed – the strategy failed when it was implemented correctly. However, it may also have been that the implementation was weak during the entrainment step. Staff weren't all on the same page, causing a different culture to emerge which led to the degradation in student behavior. Other possible outcomes could have been an incomplete cycle where the organization just stopped pursing the strategy at some point. Or, the organization could have done everything right, but observed the wrong outcomes spoiling the cycle.

Success on the other hand requires threading the needle. For a performative cycle to be successful, it must have been the basis for the design of a realistic system of practices which led staff and students to collectively live out the specified culture and, in living out the culture, affected the ultimate outcomes intended in a way that everyone recognizes.

Institutional Entrepreneurship and Strategic Culture

The most frequently studied step in the performativity cycle is the connection between theory and strategy. In the first step, institutional entrepreneurs translate theory into practicable models. In Dobbin and Jung's (2010) account, this translation process involved agency theorists promulgating new rules and polices meant to make corporate governance activities align with shareholder interests. Contrast this with MacKenzie's cases (MacKenzie 2003; MacKenzie and Millo 2003) wherein economists create financial firms to pursue trading strategies indicated by their models. Organizational theory performs through strategy – the collective agreement and formal decision making that distinguishes organizations as social actors from organizations as media for other social forces (King, Felin, and Whetten 2010).

In this study, the translation process involved the founders of the no-excuses school movement who set no-excuses culture at the heart of their educational strategy and then created charter schools to make it so. The school studied here translated this theory into a model of how instilling a specific set of values, beliefs, and norms in staff and students would lead to high student achievement. This is what I call the *strategic culture*.

I define the strategic culture as the set of beliefs, values, and norms which an organization formally endorses and expects members to abide by because the organization's leaders believe abiding by them will achieve the organization's goals. As I found, failure of the entire change process can begin in this first stage. If the strategic model is inconsistent with the theory, it is invalid. Perhaps the school succeeds, but it is not a true no-excuses school. Furthermore, if the strategy is internally contradictory or unreasonable, it is unsound.

Design and System of Practices

Design is the process by which an organization plans out the routines, procedures, and polices they will pursue to bring the culture to life. The resulting system of practices is what is commonly referred to as the socio-technical system or performative praxis (Cabantous and Gond 2011). The design process has been one of the more under-studied aspects of the performativity cycle. In their study of performativity in online advertisings, Glaser, Fiss, & Kennedy (2016) find that the practices around selling online advertisements were developed through analogical reasoning – each practice was made to look like a practice in auction markets. Performativity theorists and organizational scholars must study design and systems of practices in order to answer questions like: are theories that come with recommended practices more likely to be implemented than those that do not? Should organization develop their own systems? When do organizations create new systems and when do they adapt existing ones?

The standard for success in this stage is whether members participate in the routines, procedures, and polices. Do staff share the beliefs? Do they know the values? Do they enforce the norms? Failure occurs either by decoupling or impracticability. Decoupling is when organizations adopt a model overtly but do nothing to pursue it, or by designing an impractical system (Bromley and Powell 2012; Meyer and Rowan 1977). Schools have traditionally been decoupled organizations (Weick 1976), but this has changed with increased regulatory accountability (Timothy Hallett 2010). As I will show, the school was tightly coupled, working diligently to have staff and students produce the culture by embedding it in the schools established routines.

Entrainment and Emergent Culture

Borrowed from music and biology, entrainment describes how agents (whether particles, instruments, or people) come to resonate together across distinct activities through some control mechanism(s) (Ancona and Chong 1996; Clayton 2012; Letiche and Hagemeijer 2004; McGrath 1990). Agents can become entrained in time or in phase. Temporal entrainment is more commonly studied and involves examining how teams' or organizations' schedules, pace of work, and output align in time with others (Pérez-Nordtvedt et al. 2008). Perhaps the most apparent form of temporal synchronization in the West is shopping during the Christmas holiday.

In this case however, culture is an example of the infrequently-studied phasic entrainment – the alignment of individual, group, and organizational states. Phasic entrainment is best exemplified by national events like elections or sporting events in which large numbers of otherwise unconnected people take on the same psychological states (Fusaroli et al. 2015). The entrainment being accomplished here is to have a dispersed set of staff take on a shared set of beliefs, values, and norms by participating in many different practices. The emergent culture then is the actual achievement, the performance itself.

The standard for success in this stage is whether staff and students are able to produce and reproduce the ostensive values, norms, and beliefs in their daily practices (Bourdieu 1977; Cabantous and Gond 2011; M. S. Feldman and Pentland 2003). Mismatch can occur if entrainment fails to yield harmonization or if a different culture emerges. In his famous study of a British bank, John Weeks (2004) showed how executives' efforts to change the culture of complaint actually fed more complaint. By participating in the many programs to stem complaining, workers appropriated the programs as a way to complain about new things. The

essential empirical questions worth studying are what forces affect resonance across practices and when does entrainment lead to emergent cultures that converge or diverge from expectation?

Performative Effect and Organizational Outcomes

One of the central claims of performativity theory is that theory creates reality (Gond et al. 2016). The truth of this claim for any particular case hinges on this stage. Performative effect refers to the effects of the emergent culture. Organizational outcomes are the observations organizations make to measure the success of their strategy.

Even if organizations manage to create the culture they want, this does mean it will have the expected effects or that the effects will be recognizable. Members of the school care about things like student test scores, college acceptance, and rates of punishment. In this case, they wanted to reduce misbehavior. The performative effect they hope for is good behavior from students while the outcome they plan to observe is fewer demerits, sendouts, and suspensions. Failure at this stage means that the emergent culture failed to effect student behavior as it was measured by the school. This failure can happen for many reasons. The effect of culture could be weak. Or it could influence student behavior in ways that are not captured by the ways in which it was measured.

Narration and Cyclicality

Performativity theorists argue epistemic power lies in discourse rather than some clearly visible reality or facts (Chen 2013; Joel Gehman and Soublière 2017; Gond et al. 2016; Kornberger and Clegg 2011). We create what is real by describing it. Narration is the organizational process of review whereby members reflect on the performativity cycle in light of

the outcome, posit explanations, and plot new courses for the next cycle. However, narratives are not bounded by a single cycle. Instead, they cumulate across cycles as they repeat over time.

Performativity cycles can be made up of sub-cycles and can be part of super-cycles. In this study, the cycle I focus on corresponds with the school year. However, the school also organized reviews at different points throughout the year. These preliminary reviews were limited in their scope. Staff were more circumspect in their changes to strategy and changes to the system were limited. They were sub-cycles. At the same time, when I started my observation Achievement was already in the middle of a multi-year process of changing their strategic culture from harsh and strict to warm-strict, a super-cycle of which this particular year was only one part.

The transition from one cycle to the next, when outcomes are translated into new strategies, is perhaps the most important of all. It is here that members of the organization decide whether they believe they were successful and whether they should continue with or abandon their model. Failure in this stage occurs when narratives make the wrong diagnosis or prescribe the wrong treatment.

RESULTS

Achievement's Strategic Culture

During my time at Achievement, they were in the middle of changing their strategic culture from one based on broken windows (Wilson and Kelling 1982) to one they called warm-strict. The former combined strict rule enforcement with harsh punishments. Students were punished for the smallest missteps like getting a book from their desk ("Being Unprepared"), dropping their pencils ("Disruption"), or not sitting up straight ("Posture"). They would earn

demerits for these infractions. These demerits quickly added up to detention and suspension. For example, five years before I began my observation, half of all students at the school had been suspended at least once during the year. This broken windows-inspired strategy is common among no-excuses schools (Golann 2015; Goodman 2013).

Rather than be harsh and strict, the school's leaders decided the culture should be strict but more "warm." During the all staff orientation in 2014, the school's director Kim demonstrated their commitment to strict discipline. Citing Wilson and Kelling, she argued that strict punishment for minor violations prevented the emergence of more serious rule violations.

The problem however, as they saw it, was that harsh punishment for mild misbehavior was corrosive to student-staff relationships and hurt learning by excluding students from class. To reduce the harshness of the rules, the school changed its punishment policies by eliminating some infractions from the rule book and eliminating suspension as a form of punishment for most infractions. As a result, the school's suspension rate dropped from almost 50% in 2010 to 10% in 2013.

To make their culture more "warm," they advocated a set of mitigation practices teachers should use when managing their classroom including "proactive management," "rational detachment," and "assuming the best." Using these practices, teachers could prevent rule violations and enforce rules without making enforcement personal.

The strategic culture then was meant to remain strict. No student would get away with the slightest misbehavior. But, the approach to enforcement was warm. The methods of enforcement would be explained ahead of time. Classes would be designed to support abiding by the rules. Violations by students and enforcement by teachers would not be addressed as personal affronts. Teachers would approach students' motives with the presumption of good faith. The result of

this strategic culture would be well-behaved students and cohesive relationships between students and staff.

As I said, every stage has its own form of success and failure. In this case, the warm-strict strategy certainly seemed consistent with a culture of high expectations. All rules would be strictly enforced for all students so that students would eventually learn to be well-behaved.

However, there was an internal conflict in the strategy. The presumption of good faith, "assuming the best," gave teachers a degree of discretion that meant not all rule violations would be punished. This conflict between discretion and strictness begs several questions. First, if teachers are letting some students off the hook because they did not mean to break the rules are they simply making excuses for the students? Is discretion another word for excuse? If this is the case and they succeed in reducing misbehavior, will they still be able to say it was the no-excuses culture that led to their success? Or would success be the result of abandoning their culture? Alternatively, if they fail, will it be because they abandoned no-excuses or because no-excuses theory really doesn't create a good culture or because they were implementing a weak model?

How the school answers these questions matters. If they are successful by mistake, it would seem less likely they could replicate or sustain their success next year. If they are successful for the wrong reasons, they may incorrectly double down on warm-strict (a false positive) and future strategies would undo their success. If they fail because their model was flawed, they may prematurely reject warm-strict altogether (a false negative). In the performative model of change, we want to be successful because we knew what we were doing and did what we said we were going to do. If we are unclear about what we are doing, we cannot distinguish

true and false positives or true and false negatives. Without clear and consistent models, we cannot learn from success or failure.

System Design at Achievement

To create a warm-strict culture, the school's leaders created the "culture of consistency" initiative for my focal year. The plan was to get teachers on board with the warm strict model and have them use it to establish consistent discipline across all classes early in the year. They believed this consistency would create a culture of good behavior that, once established, would be self-sustaining.

Rather than create new trainings or classwork, their plan was to repurpose existing routines. They would train teachers during orientation and use the first two months' worth of regularly scheduled faculty meetings, department meetings, and grade level team meetings to help teachers align their discipline with one another. Once this initial push was over, staff would return to the normal work accomplished in these routines.

They executed on this plan. They presented warm strict culture to teachers during orientation. Department chairs and grade level team leaders set aside time in their meetings to discuss classroom culture and consistency. An analysis of the frequency of the root "consisten-" in meeting agenda and notes from grade level and department meetings (Figure 3.2) shows that consistency played a substantial role in the first two months of meetings and then declined precipitously through the fall and winter until it spikes again in the spring.

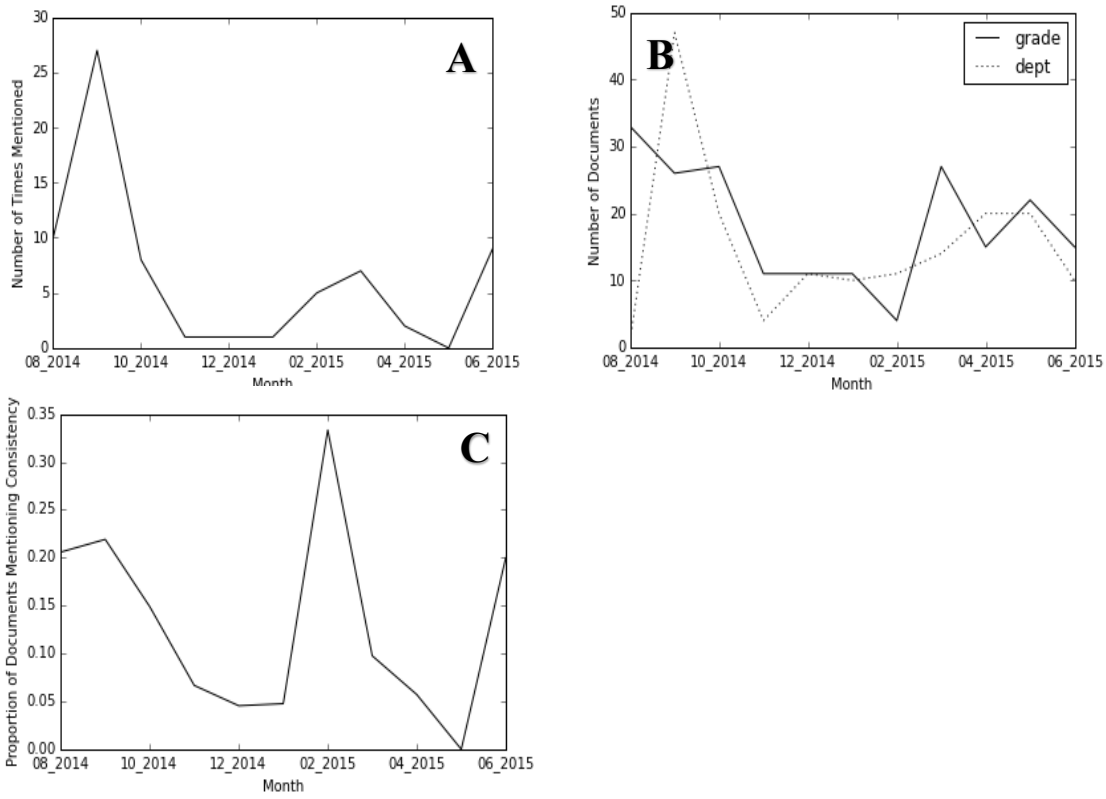

Figure 3.2: Frequency of Meetings and Mentions of Consistency in 2014-2015. A) Monthly Number of Times "Consisten-" Mentioned in Grade Level and Department Meeting Notes, B) Monthly Number of Meeting Notes from Grade Level and Department Meetings; C) Monthly Proportion of Meeting Notes mentioning "Consisten-"

Success and failure in design and systems are dependent on the clarity of the design and practicality of the system. The design was very clear to teachers. Achievement engages in various initiatives each year. Further, the idea of establishing culture early is common and well-known in education. If you tell teachers they need to establish rules and routines early, they will understand what that entails and why it should work. Most have heard it before and pick up on it quickly as a normal thing to do. In this way, the design succeeds in being simple and normal rather than convoluted or strange.

The simplicity may also have been a problem. It was so easy to envision the strategy working that they assumed it would. When the culture began to unravel over the winter, they had no plan to compensate and scrambled to make the correction.

The system was practicable. Staff were already supposed to be at orientation and to participate in faculty meetings and grade level teams. School leaders would not have to worry about scheduling new events, communicating with teachers, or justifying yet another meeting. The time is already set aside and the people planned to be there ready to work. On the other hand, as I show next, putting the culture of consistency initiative into extant routines led them to bleed into what teachers normally do in these meeting and what they were asked to do as part of the culture of consistency initiative. This prevented teachers from entraining on a consistent approach.

Emergent Culture

There is a difference between what you want to do, what you actually do, and what you ultimately accomplish. Emergent culture represents that second step. At Achievement College Prep, emergent culture is made up of the beliefs made salient, norms established, and values valorized. In my focal year, three types of action constituted the emergent culture: the realized practices, ancillary actions, and exogenous actions.

Realized Practices

The school followed through on its plan to discuss consistency in its normal routines. I refer to these actual conversations that took place according to the plan as realized practices. The effect of these meetings was entrainment: getting staff "on the same page" around consistency. However, the way consistency was discussed was not always congruent with the stated strategy.

> Orientation was a forum for administrators to explain the strategy to staff, provide staff with scripts to justify the warm strict model to one another and students, and have staff practice disciplining students through the new model. In the following excerpt from new staff orientation, staff were put into groups and given disciplinary scenarios. They had to determine if a student's actions warranted a verbal warning, demerit, or send out and then role-play giving the punishment to the student:
> The scenario on the PowerPoint reads: "In class, Grover is slouched down even after a whole class reminder about posture. You say, 'Grover, you have a demerit for posture' and he responds, 'for what?'"
>
> The consensus is red cards for a send out. (Students are not allowed to react to a demerit.) One teacher who didn't give a red card speaks up, "Maybe they didn't hear?" "No" David says "'For what' is usually not an 'I didn't hear,' but an asking for justification or a complaint."
>
> To one of the teachers who held up a red card, Rich says: "Okay, give me the demerit." The teacher obliges, "Rich, you have a demerit." "FOR WHAT?!" Rich responds heatedly. The new teacher is taken aback by the character Rich put on and stumbles into silence. David picks up the conversation, "what's the terminology for back talk?" Another teacher reiterates, "But what do you do if they don't hear you?" "They'll hear you" David answers and then demonstrates the recommended response if they don't hear, "you have a demerit for posture, sit up."

This excerpt exemplifies how orientation provided teachers with the categories, beliefs, and norms they would need in the classroom. There was practice classifying behaviors – "students don't say 'for what?' because they don't hear you." When responding to a student, the tone was to be calm and direct: "you have a demerit for posture, sit up." During orientation, teachers developed the emotional, technical, and normative skills needed to do warm strict discipline in preparation for creating a culture of consistency.

Once the year began, teachers' discussions of consistency changed to be problems of instruction and systems, reflecting the concerns of department and grade level teams respectively. In general, department meetings were meant to help teachers develop their curriculum (what is taught) and instruction (how it is taught). In department meetings, the culture of consistency initiative was treated as an extension of instruction.

During the first math department meeting of the year, the chair, Matt, went through the culture issues in his class: "I was trying to have a 25 person discussion at 3:00 when the room was 92 degrees. I projected poorly. I was too quick to call on someone. I was like, 'I got three more things to get to. You!'" Discipline is not the only way to engender a set of values, norms, and beliefs in classrooms. Culture is just as much a result of the activities students do in class, who a teacher calls on, and how mistakes are managed. In focusing on creating culture through instruction, departments translated the initiative into a different and incommensurate frame.

In contrast, one of the primary roles of grade level teams was to ensure teachers in a particular grade level enforced the rules consistently. The culture of consistency initiative should have fit directly into the normal conversations of grade level teams. This was not the case, however. In the following note from the twelfth-grade team's first meeting of the year, teachers report many divergent practices and the conversation does not make them consistent:

> Dawn, the grade level team leader, started the meeting by having teachers write down notes for anything they wanted to discuss. The teachers write silently and then Dawn brings them back, pointing to Maggie, one of the teachers, "I want to start with you." Maggie begins, "Seniors are still having trouble reading the schedule. They're not used to it. I still had to tell them, 'hey buddy, you're up here [in this room].'"
>
> Edith adds, "I've been using more send outs." Dawn replies, "We're trying to work on other teacher tricks before it's a send out." "You can only try so many teacher tricks," Edith says, "For my first class, only six students were there for the first half because they were in college prep." Another teacher seconds this, "I'm seeing late seniors in the hallway. Just, between classes, I have a few tardies." Matt says, "I want to ditto what's been said. I'm okay not doing them (giving tardy demerits to seniors). But I've not been told, 'don't do it.'" Dawn says, "I think it goes along with what you want to do. What I've heard from seniors is that they like that things are different from one teacher to the next."

This note reveals three ways in which teachers' actions in the grade level team diverged from the culture of consistency initiative. Whereas the consistency initiative sought to reduce

punishment by creating a consistent culture, both Edith and Dawn reverse the causal arrow using punishment to create culture. This may seem like splitting hairs, but it was an article of faith among the school's leaders that, "you can't discipline your way to good culture."

Second, there was a counterculture among twelfth grade teachers who viewed inconsistency as more "collegiate" because college courses all have different rules and expectations. More generally, there was a pattern in the school of becoming more lenient on students as they advanced to higher grades. Finally, in this meeting and others, teachers uniformly reported being lenient on tardy demerits early in the year because of students' schedules. Most notably, the meeting does not conclude with a plan to create consistency, but instead to institute it.

Conversations early in the year indicate the ways in which teachers' behavior converged and diverged from the culture consistency initiative. Even though they followed through on the system, the actual discussions were not faithful to the logic of the initiative clearly, uniformly, or cogently.

These divergent practices ultimately undermined teachers' sense that they were consistent with one another. In interviews and observations of end-of-year review meetings, staff I talked to reported that the school never "felt consistent." For example, one teacher reported there were things she would do during transitions that other teachers wouldn't and that, during grade level meetings, they would discuss consistency but never actually followed through on making any changes. Another reported that a colleague "just decided to do whatever the hell she wanted because that's just how she is."

Exogenous Actions

Every stage in the performance cycle is susceptible to disruption from outside forces. During my focal year there were two major disruptions. First, the school continued to change its systems and procedures throughout the year, making it difficult for teachers to keep up with what rules to apply when. Second, students surprised the administration by participating in a walk out to join a Black Lives Matter protest, compelling school staff to address the role of race in the school. I address each in turn.

The school is built to change. Grade level teams, the policies and procedures committee, and high school and middle school teams are all tasked with tweaking systems and procedures and all meet on a routine basis. As they changed, they undermined staff's ability to know what rules applied when. In the following field note, the policy and procedures committee, made up of grade level leaders and senior administrators, discuss changes to Community Meetings:

> Kim continues down the agenda, "Someone mentioned transitions to community meetings?"
>
> A teacher begins, "So two parts: theme of silent transitions. That falls to individual teachers. The other thing is getting students out [of the meeting]. The idea is to send 6th and 8th out simultaneously using separate staircases. We want to exit four homerooms at a time."
>
> "It's a traffic jam" another teacher adds. The teachers discuss among themselves the potential of moving different groups of 6th and 8th graders down from the main area to the garden level without creating a big mess of people.
>
> The director summarizes, "So one eight and one sixth grade class, then the second eight and sixth? We just have to learn how to say that [as an announcement at the end of the meeting telling who to get up when]. Two teachers have a side conversation among themselves, clarifying which classes to go when. They come up with a solution and Kim says, "We'll have to clarify that" [clear up the language].

What is most important for understanding this conversation is that the school has been doing community meetings every week in this building for almost ten years. After a decade, they are still tweaking even the most basic aspects of one of the longest-running procedures in the school. These constant changes make it difficult for staff to keep up with what rules they're supposed to be enforcing.

The second disruption occurred right before winter break when a number of students walked out of school to join a local Black Lives Matter protest. The walkout sparked conversations about the role of race in the school. Students told staff they felt the school was "white space" and that the mostly white teachers did not understand their experiences as black people.

In a series of meetings starting in January and going through June, staff discussed race and school culture including topics like the role of race in hiring, school discipline, and relationships with students. Staff explored their own implicit cognitive biases through conversations and by taking the Implicit Association Task for race online (www.projectimplicit.com). They gathered feedback from other schools, their board, and community organizations on what they could do to be more inclusive. They put out a survey to staff and students to elicit people's interests and hobbies to match students and staff with similar interests in the hopes that they would build relationships on them.

Throughout the second half of the year, the staff at Achievement poured a lot of time and energy into re-thinking the culture of the school through the lens of race. During these conversations, staff did not directly engage warm-strict or the culture of consistency initiative. Instead, the discussions took time and energy away from the culture of consistency initiative.

The culture that emerged over the spring, was a more race-conscious, reflective, introspective, and vulnerable one than imagined in their consistency initiative.

Organizations are open, complex systems. Performance cycles are never isolated from other parts of the organization nor are they pursed in a cultural vacuum. In fact, every stage in the performativity cycle is susceptible to forces beyond its scope. For Achievement during my focal year, the culture that emerged was substantially different from the culture that was planned in part because of ancillary and exogenous actions taken by staff and students.

Culture, Good and Bad

In the ideal case, teachers would have learned to balance discretion and consistency, worked with one another to maintain that balance as new issues arose, and overtly appealed to the tenets of warm-strict culture – "reading intent," "proactive management," "rational detachment," "assuming the best" – in addition to the school's other values and beliefs. It is not unreasonable to imagine this culture emerging. In fact, they did come back to the culture of consistency initiative after the initial push early in the year. And throughout the school year, teachers understood the consistency initiative well enough to discuss it on their own.

At Achievement, the culture that emerged deviated from the model in three ways. First, the beliefs and norms conflicted and blended with those embedded and defended in other routines and groups. The school's constant changing of its rules and systems undermined the basis for consistence. And another culture displaced the initiative as teachers focused on a new set of beliefs, norms, and values. All three of these deviations could explain the ultimate failure of the initiative. Yet, had one not happened, who knows whether or not the other two would have been enough to undermine the culture in its own right. A good model executed poorly performs as well as a poor model executed well.

Performative Effects and Outcomes

In the ideal case, a well-defined cultural strategy is practically designed and carried out, resulting in it emerging in the daily life of the school. If the model is correct, the emergent culture should then cause the desired change in organizational outcomes to become visible. This step from achieving a certain culture to producing specific organizational outputs is the performative effect.

At Achievement, the culture of consistency initiative should have established a self-sustaining warm-strict culture early in the year, leading to better student behavior and a decline in disciplinary actions across the school year. However, the fact that the culture that emerged was inconsistent created the space for a lack of discipline. First, inconsistency gave students the means to resist discipline leading to further inconsistency and consistent leniency. Second, inconsistency discipline eroded trust among faculty and students. These two effects ultimately drove an increase in the school's observed outcome – disciplinary actions in class.

Performative Effects

The culture of inconsistency gave students leverage over staff. Because the school's authority system was built on consistency everyone knows the rules. Thus, if a student responds to a demerit by saying, "so-and-so doesn't give me a demerit for that," teachers are put in a bind. They can either give the demerit and undermine the other teacher or not give the demerit and be inconsistent with the rule. Often teachers would decide to give a verbal warning rather than punish the student. Inconsistency bred inconsistency.

Some students would use this to get permission to break the rules. For example, one seventh grade student managed to get one teacher to allow him to sit sideways in his chair during class, a violation of the posture rule. The student then sat sideways in two other classes, telling

the teachers that the first allowed him to do it. One teacher resisted the ploy, but the other allowed it. In this way, inconsistency also bred consistent leniency.

The culture of inconsistency also eroded the trust teachers had in one another and in their students. Again, because everyone knows the rules, it is obvious when a teacher is not enforcing them. The following discussion from a meeting of department chairs in April exemplifies this:

> Ellie says that the substitute was talking with students in the middle school cafeteria when the students should have been going back to class. Another teacher wanted to correct the students but didn't want to undermine the substitute's authority. "The students were with an adult" is how Ellie phrased what the teacher had said. Ellie concluded saying, "There's a felt perception that she's not competent."

In this scenario, the substitute is seen to be openly breaking the rules with students by talking with them in the cafeteria during class where anyone can see them. The point for Ellie in telling the story is to communicate how untrustworthy the substitute is to the other teachers. More generally, trust did not erode among staff as a whole but for particular staff who were seen as not toeing the line.

The mistrust among staff was complemented by a mistrust staff developed of students. Returning to Figure 3.2, the spike in mentions of consistency in February and March stem almost wholly from middle school grade level meetings. In February, middle school teachers began discussing consistency again. By March, consistency was being discussed as a systemic problem. According to notes from their March 9th meeting, the sixth-grade team began implementing "resets" – halts on all classroom activity until students settle down. And on March 10th, the eighth-grade team discussed consistency in terms of how to get their students invested in the class. On March 18th, the seventh-grade team brainstormed ideas for how to address issues of collective misbehavior.

The lack of trust bubbled over for the eighth-grade team in May when they discussed not going on their annual overnight field trip to visit a college. In the following field note, the grade level team is discussing ways to mitigate the dangers they see in their students:

> When I come in, the room is cold and quiet. They're discussing whether to hire a security guard to monitor the hallways of the hotel during the 8th grade's overnight trip. I'm immediately struck by the gravity in the room and wait to take notes to see where the conversation is. The teachers are very hesitant about certain students who they believe could get themselves into trouble. At one point, a fourth-year math teacher Danielle acknowledges, 'we've had rough classes before, but not as rough as this class.' That encapsulates the focal point of the conversation.
>
> Scott a new art teacher says, "If the likelihood that something happens is 5%, then if we do this for so many years, something is going to happen. We need a plan for when it happens." Kim adds to this, "Yeah, on the Rome trip, we've had a fight, drinking, smoking. Stuff has happened and people on the team have managed it." Seeing no responses in the room, Kim asks bluntly, "Would you prefer it if it's not an overnight. Do you even want to go?" Justin says with quiet conviction, "I want to go and I think it should be overnight. The kids like it. And, it's a rite of passage and we don't have many of those as Achievement."

The director walked out of the meeting furious. The idea of hiring a security guard or cancelling the overnight trip altogether demonstrated how eroded trust had become and flew in the face of the warm-strict idea of "assuming the best." But teachers thought the kids were particularly dangerous. So dangerous in fact that they considered cancelling what most had come to see as a rite of passage.

The school's cultural inconsistency had the effect of eroding trust and creating inconsistent or consistently lenient discipline. These effects however are different from the outcomes observed by the school. For school leaders, the theory was that consistency would reduce misbehavior. The reduced misbehavior would be observable in the number of demerits, send outs, and suspensions. These numbers were ultimately the outcome of interest to the school to the extent that they presented evidence of good behavior.

Outcomes

Outcome measurement (and review) began as early as the first faculty meeting of the year. Here is the Dean of Student Life presenting the data at the faculty meeting in early September.

> Rich began, "So, send-outs and demerits are up over last year. They're up 50% [663-977] and send-outs are up 10% [70-76]. We're going to go into what students are getting demerits and send-outs for. It's important for us to keep in mind our goals and what we're trying to accomplish. So, categories for demerits. What you're going to see here is, areas having the highest growth in demerits. So, food is definitely one, off-task, speaking out of turn, tardy. And this is school wide."
>
> (the teachers discuss the numbers at their table)
>
> After teachers shared their thoughts, David said, "We definitely don't want you to think that, because the numbers are up, that those are bad things. This is not necessarily a bad thing. We have to do this so we can correct things as the year goes on."

From the beginning of the year, the rate of discipline had skyrocketed. But they did not automatically interpret it as a sign of failure. The idea was to bring down the rate of misbehavior as the culture of consistency initiative continued and the warm/strict culture became established. However, the default response from staff is still that the elevated numbers are not good. That's why David says the numbers are "not necessarily a bad thing."

The numbers remained elevated throughout the year, especially for middle school. Figure 3.3 compares the rate of send outs during the 2014-2015 school year for the middle school to the prior year's send outs. In all, the number of send outs in the middle school doubled from 2,000 in the 2013-2014 school year to 4,000 the following year. Only the seventh grade saw a year-over-year rate of growth that was less than double. Sixth grade saw an increase of 150% from 500 to 1250 while the increase for eighth grade was over 200% from 350 send outs in a year to 1100.

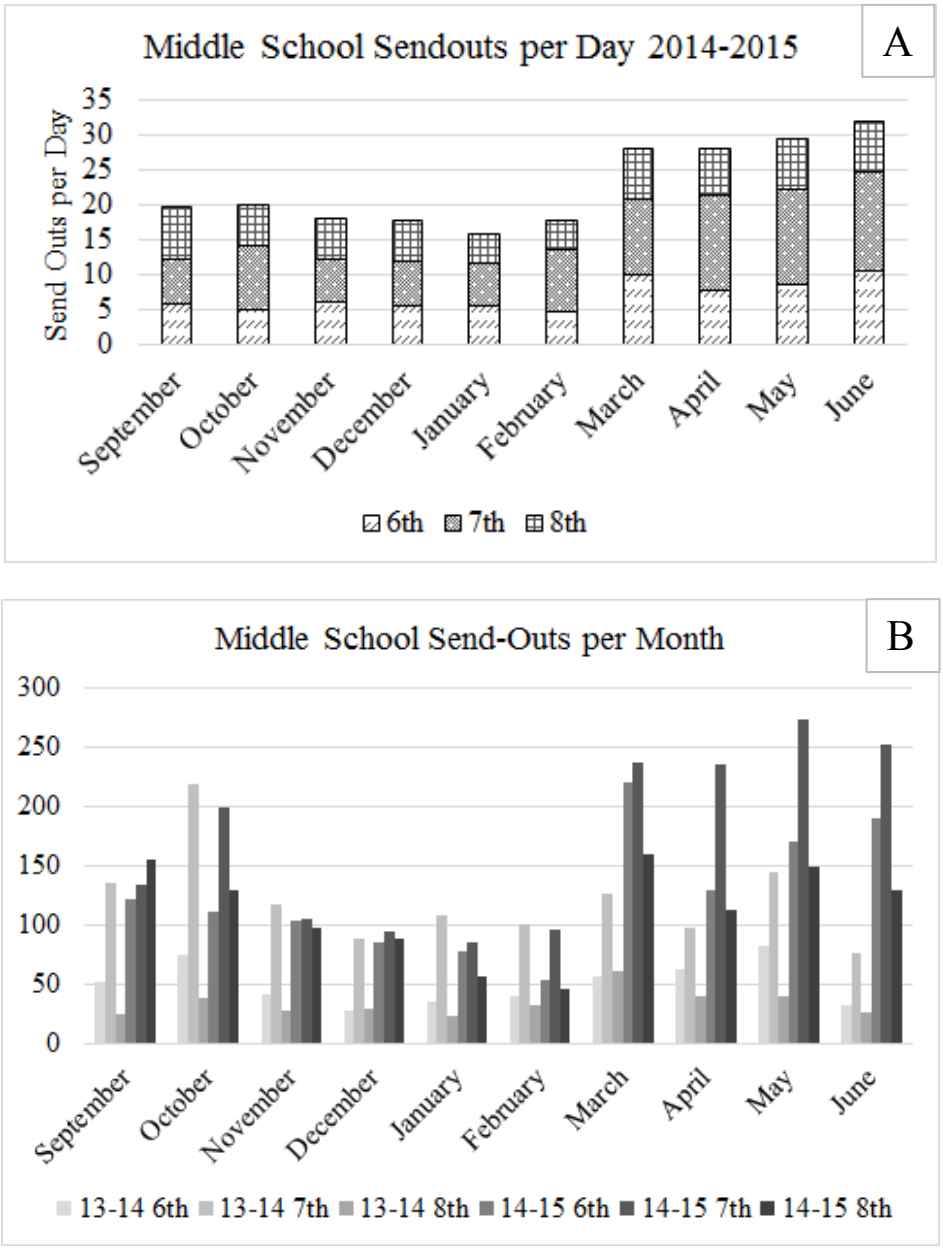

Figure 3.3 Distribution of Send-outs in the Middle School (A) Per Day in School During the 2014-2015 School Year and (B) During the 2013-2014 and 2014-2015 School Year.

In Figure 3.3A, I plot the send outs with the percentage of days the school was in session. The send outs per day shows that the decrease from October through February can be modeled by the days in session. It also throws into relief how extreme the increase in March was. There were 50% more send outs in March than October despite their having the same number of days

in session. Similarly, the number of send outs in April was two thirds higher than in November even though both had 17 days of class. Seasonality substantially affected the number of send outs across the year, but it did not determine the underlying *rate* of send outs.

The disciplinary data reveal three changes occurring during the 2014-2015 school year. The first was an increase in discipline across the entire 2014-2015 year as compared to the previous year. The second was a doubling of discipline in September 2014 compared to September 2013. The third was a extreme increase in March 2014. What explains these changes?

Inconsistency and consistent leniency do not explain these patterns in the data as both suppress the number of disciplinary actions. To be lenient is to not give demerits. Only one thing changed year over year that also changed in September 2014 and then again in March for the middle school: staff's focus on of the culture of consistency initiative. the "culture of consistency" initiative had a framing effect on staff whereby the driver of classroom misbehavior was a lack of enforcement.

As a frame for understanding how to discipline students, "consistency" encouraged faculty to see culture through the lens of punishment: am I punishing everyone the same? Returning to the first department and grade level team meetings of the year, the impact of the reversal in logic becomes clear. Staff saw the initiative as setting punishment as the driver of good behavior and not punishing as the target outcome. For all the potential challenges created during the performativity cycle – navigating discretion and consistency, communicating consistency, the loss of trust and consistent leniency; the best predictor of discipline is staff using the consistency frame to diagnose problems with classroom culture.

Why did they Fail?

Failure is measured in two different ways. The first is conceptual – did we achieve the impact we sought? In this case, did they create a warm-strict culture and increase trust? They didn't think so. The second is the explicit measure – did they reduce the number of disciplinary incidents? On this they failed as well. In one respect, they failed because they didn't achieve what they set out to achieve. In another respect, they failed because the outcomes they were measuring told them they failed. However the modes of failure are distinct.

Their failure to create a warm-strict culture and instill trust actually suppressed the number of disciplinary actions. Lenient teachers give out fewer demerits. The increase in disciplinary actions throughout the year is a result of the consistency initiative itself. The initiative encouraged staff to think about classroom culture through the lens of discipline as opposed to other mechanisms like proactive management or working with parents. If there were disciplinary issues in class, then they were the result of students getting away with stuff and needing to be punished anew.

Review

While I have pointed to many of the weakness of the culture of consistency initiative and warm-strict strategy, these remain my own. What matters to the members of the school and to the field of education is how the staff interpreted their successes and failures. These interpretations are the basis for whether and in what way the school remained committed to the warm-strict strategy (and ultimately the no-excuses project).

Narrating Change

As soon as the year started and outcomes became observable, the school began to review them. They reviewed disciplinary data at the beginning of the year, again in spring, and finally during a series of meetings at the end of the academic year. During each review, staff came up with narratives of what happened and tweaked their strategy. Each review represented a performativity sub-cycle and the end-of-year review marked the end of the larger, annual cycle which itself was part of a larger super-cycle.

Outcome reviews were so common at Achievement that there was a pre-defined agenda for them. Whether they were grades, attendance, test scores, or demerits; school leaders would give staff summaries of data in graphs, tables, and charts; had staff discuss the data in small groups, and then collectively discuss solutions. These conversations were always guided by the same two scripted questions: "What does the data tell you?" and "What are the implications for the school?"

Review began with teachers interpreting the data and offering explanations in light of their own experiences and beliefs. At the data review meeting in September, some teachers theorized that perhaps the elevated rates of discipline were due to students who started the year late. Another suggested it may be because there were more students in the school that year. During the data review in May when discipline continued to be high, teachers offered different theories. One said the dean's office wasn't a punishment anymore, students were acting out to go to the dean's office. Another suggested that it may be that they had more first year teachers. Another suggested there weren't enough adults monitoring students.

By design, these interpretations were followed by change recommendations. In September, the ninth-grade team thought they were being too lenient and decided to discipline

students more rigidly. To prevent tardy demerits, the tenth-grade team suggested letting students out on time so they have time to get to their next class. The twelfth suggested they focus on demerits like tardiness because being late to class would hurt students in college. By May, teachers suggested more general strategies like "creating a culture of responsibility," having formal conversations with students to prepare them for the next grade, and making classes more engaging.

The number of interpretations and ideas increased as the year concluded with the annual end-of-year review. Staff participated in grade level and department meetings in which they reviewed their annual goals and discussed the things they should do differently. Staff were also put into teams diversified by tenure, department, and grade to discuss what they should do differently for culture in addition to curriculum, college readiness, and school operations.

Across these sub-cycles in September, May, and June; staff turned the many interpretations and tweaks into narratives of "what happened." These narratives informed the bevy of changes to the school's systems during the year and to changes in strategy from one year to the next. This is one "what happened" narrative from one senior administrator given in an interview at the beginning of the next school year.

> For students, I think the culture has shifted just because we have more stuff for students to do. So that creates its own shifts. That's both clubs, athletics, academics. I think we're seeing increased pushback on our systems and I'm wondering if the pendulum is swinging a little. My understanding, when the school started it was like, boom-boom, very rigid, very this this this. They would get into like battles with kids and kids would leave over the *extreme* nature of some of the rules. I think we've swung over to a place in the middle where people are have an understanding of it, getting it, kind of bought into it, and now we're starting to see more pushback on some of those rules and systems which is impacting culture.

This narrative sets experiences in the 2014-2015 school year in the context of larger-scale narratives from prior years in which the school is trying to engage students rather than control them. This administrator's broad diagnosis is that, now that students have learned the warm-strict system, they are starting to resist it. The phrase this administrator uses which many others use is the notion of that the school's strategy is a pendulum that has swung too far towards leniency.

Given the plethora of suggestions and failure to curb misbehavior, how did the school change its strategy at the end of the year? First, the school's leaders doubled down on culture in general and warm strict culture specifically as a focus for the school. For example, the amount of time they devoted to culture during new teacher orientation grew from three and a half hours in 2014 to eleven and half hours in 2015 (it was only an hour in 2009).

Second, a new initiative was undertaken in 2015 to "create a culture of opting in." This initiative focused on creating warm-strict culture by getting students to engage in class as a means for achieving academic and behavioral goals. In addition to the new initiative, they retained the system established for the culture of consistency initiative. That is, they discussed consistency in meetings early in the year and then planned to move on. However, they would move on only after they felt consistency had been achieved.

Finally, they continued to make substantial changes to the most basic systems and procedures. For example, the time set aside for the ethics curriculum in middle school, a cornerstone of the school's mission since its founding, was moved from homeroom to lunch. These changes were proffered with the caveat that teachers should only change them rarely so everyone can stay consistent.

Performativity Cycles

Given the poor results and so much review, why did the school not abandon warm-strict, consistency, or even the no-excuses theory? I believe it is because staff narratives were fundamentally conservative in several ways. First, staff resisted attributing causes to students' behavior. As one administrator said in an interview, "There was something about the egregious behavior [in the middle school]. It wasn't shut down by other students as much as it had been in the past. I'm unclear as to why that was and [we're] really trying to be mindful about some changes to make that not occur." Having observed the system for a year does not necessarily mean that staff believed they understood it. Many were unwilling to make direct claims to know why things happened the way they did.

Second, these narratives were also trapped in time. Many things happened in the 2014-2015 school year that will never happen again. Due to miscommunication, they did not have a dean of students for the first month of school. An unseasonably snowy winter led to an unprecedented number of snow days in January and February. The highly publicized shooting of black men and women by police officers across the country traumatized students. In evaluating the consistency initiative, staff recognized they only had one year's worth of data on which to evaluate the initiative.

If such gross failure during a given year does not change the strategic culture, what does? Looking to the school's history, the change to warm strict itself represented a substantial change in their strategic culture away from the classic no-excuses model of harsh and strict discipline. The school began making the shift away from harsh and strict punishment when they hired a new executive director three years earlier. Archival documents and conversations about this transition with staff indicate that the prior director was directly involved in finding the new director and

wanted to make the change away from harsh and strict punishment. The will to make the strategy change was there. Hiring a new director was a way of further encouraging the change. This suggests that what led to that shift in strategic culture was a performation super-cycle with a longer time horizon and broader scope. What can be changed at the end of a cycle is relative to the scope of the cycle itself.

DISCUSSION

Over the past four decades, we have developed a wealth of scientific knowledge about organizational culture and the benefits good culture has on organizational performance, especially for schools. We should expect good schools to have good cultures and effective schools should be able to create them effectively. This was not the case with Achievement during my year there for proximal and distal reasons.

At each step in the cycle, the school did or experienced something that undermined their attempt to reduce student discipline by creating a culture of consistency. The way the initiative was framed as being about disciplinary consistency led staff to address misbehavior through punishment, driving up their rates of discipline. By appropriating existing routines rather than creating new ones, staff blended the logic of the initiative with those already established in their local routines, oftentimes in improper and inconsistent ways. Unexpected events like the Black Lives Matter protest displaced the emphasis on consistency at a time when it was most needed. It is difficult to say that fixing any one of these issues would have led to success. Each surely contributed to the failure of the initiative in its own way.

Ultimately, the most important reason they failed was because the road to success for performative change is razor thin. When organizations have the clarity to create their own plan

and the operational control to implement it, the points of failure multiply and become obvious. With control, organizations internalize all of the risks and uncertainties of implementation. Every step in the performativity cycle is a possible point of failure for which the organization is ultimately responsible. Clarity entails that everyone in the organization, to one extent or another, knows what is supposed to happen in each step and can tell when a step is unsuccessful or not.

There is another distal, yet more essential, reason Achievement failed: there was no clear model for success. There was no proven culture program for them to implement. Though we have plenty of ideas for what good cultures should look like there are no programs schools can adopt to create these cultures. It is one thing to say "you should reduce the salt in your diet" and another to say "you should create a diet plan to reduce the amount of salt in your diet." In education and organizational sciences, we believe we know what kinds of cultural milestones an organization should achieve, but we have nothing to say about how best to get there. So it is that no-excuses schools offers a set of beliefs, norms, and values according to organizational theory; but theory offers little in the way of how to go about using these to achieve strategy goals. As a result, the way in which these schools go about using culture differs substantially and idiosyncratically.

This lack of prescribed programs makes it impossible for researchers or practitioners to assess the effectiveness of a cultural strategy. At the end of the year, staff at Achievement had a strong sense that the strategy had failed – the pendulum had swung too far towards leniency. But they differed in assigning blame to the logic of the strategy, the design, or staff's execution. Without other points of comparison, whether other schools or experience in prior years; there was no way for them to know what exactly went wrong. Even for scientists, we have putatively

good theories of school culture but no way of knowing whether they emerge in schools on purpose or by accident because there are no standard implementation programs.

Finally, the school failed because they were working with science that is thirty years out of date. The no-excuses culture at the heart of Achievement was invented in the 1980s, used to build charter schools in the 1990s and 2000s, and has only been producing outcome data for the past ten years. By the time I began my observations, the theory at the heart of Achievement was three decades old. This does not mean that the theory has been proven wrong. I am simply pointing to the lag in the time in which a cultural strategy is invented and when we see schools trying to use that strategy. In this case, the no-excuses culture diffused among the population of schools not through the communication of science-based best practices, but through population replacement driven by a social movement. This ecological approach to implementing science-based practices is slow and epistemically inefficient.

Not only is it incredibly difficult to successfully pursue performative change, but there are few supports for attempting performative culture change. And, the study of culture has not lent itself to proposing interventions that work and schools are not built to change cultures with the latest science. If we want to see more successful performative changes in schools, we need to provide proven implementation models.

CONCLUSION

Ultimately, scientists, policy makers, educators, and organizational leaders all want performative strategies. We want to know what we can do to solve our problems and we want a proven set of steps to implement the solution. Performativity theory provides a general model for how research can provide such a solution. One hopes that by better theorizing performativity and

studying its cases, we can develop successful theories with positive effects. From this case, it is apparent that our decades of research on organizational and school culture have amounted to very little in the way of a system of interventions for improving schools in a performative strategy. Even for putatively effective schools, creating effective culture is a local, idiosyncratic, and ultimately isolated affair.

There is however a well-developed model for developing effective performative strategies: medicine. In medicine, research leads to the development of interventions like medications, surgeries, or therapies with defined the treatment plans which work at some predictable rate. Medicine is not a perfect field where all problems are solved and all solutions work. It has well-established norms, processes, and institutions supporting the development, evaluation, and certification of performative interventions. We are far from this model in organizational science and education. We are more like the field of nutrition in which scientific findings get picked up by entrepreneurs and marketed through best-selling books but are rarely rigorously evaluated until after the fad has died out.

Rather than using randomized control trials, the performance evaluation that does occur is post-hoc, examining those schools or organizations that happened to adopt the intervention we care about. At Achievement, it is impossible to know what errors really led to their failure to achieve a warm-strict culture without systematic comparisons to other schools trying the same strategies but perhaps making different errors. Because change is ad hoc and idiosyncratic, there are no comparable schools undergoing comparable culture change. At the end of the year, not only do they not have clear evidence for whether their strategy worked, but neither do researchers.

CHAPTER 4

COUPLED TO THE CORE, EDUCATION'S MARSHMALLOW CHALLENGE

INTRODUCTION

It was one of my last observations – a faculty meeting one Monday night in November 2015. The tricks played on my circadian rhythm by the early onset of night that accompanies the Fall made the otherwise mundane faculty meeting feel tiresome and dreary. The meeting had mainly been about the schools' new culture initiative and how the culture among the adults was going. But after a brief update about faculty's upcoming visits to other schools, the meeting took a sharp turn – the schools' test scores had come in.

> The director introduced the topic: "So we want to share our initial results as a school. The results have [already] been shared with the Math and English teams. It's mostly their data, but it's not just theirs. The results are not where we want them to be. It's not where the Math department wants them to be. They're not where English wants them to be. But, we can agree that we think the results are honest. It's where we think we are. It's a higher bar. Math and English are out in front of the response."
>
> Kelly, the new Director of Academics, shows the results from English and then from Math. They follow earlier patterns where the school underperforms in earlier grades, but in both Math and English perform above average in the latest grade tested (9th grade in this case). There is a green highlight for the scores where the school performs above the state's average. This makes it easy to see the comparison. However, the student growth numbers have no point of comparison and so are difficult to interpret. They range from around 30 to 60.
>
> Kelly then takes questions from the staff. She explains the new scale: the state's traditional qualitative levels of passing, needs improvement, failing, and so on have been replaced with a 5-point scale in which 4s and 5s are satisfactory and 1s, 2s, and 3s are underperforming. One teacher asks how they're calculating student growth (one of the scores shown on the PowerPoint) given there was no PARCC last year. Kim answers that they're using a formula comparing prior scores on the state test to current PARCC scores to create a cross-test growth score.

> Kelly concludes, "So there are next steps from here. There will be more data released this week. We know we need to respond as a school to this. These are all our results."

The school's scores were mediocre and met with widespread disappointment. Until PARCC, all of their students had been passing the state's test by the time they graduated from high school. Now, fewer than half of the school's high school students were considered proficient. For the middle school, only twenty or thirty percent of students met that goal.

Organizational and educational scholars tell us that teachers and administrators should have rejected the test particularly because, at the time, the state had already decided to scrap the PARCC following widespread criticism of the test. They should also have rejected the Common Core as a superficial policy with little practical relationship to the demands of their classroom. In fact, education has long been the poster child for decoupling (Coburn 2004; Deal and Celotti 1980; Weick 1976). Schools, administrators, and teachers are supposed to be intransigent to one another and the state.

Achievement College Prep didn't decouple. They did not challenge the test or the standards. Instead, they embraced them. They thought the test was a reasonably accurate reflection of their students' learning. They believed their students did not know enough to pass the more rigorous PARCC exam. Their failure served to deepen their coupling. They doubled down. They would work even harder to be held accountable.

However, the experience of Achievement College Prep points to a fatal flaw in the school accountability regime. It cannot lead to more effective schools in a reasonable time horizon. The logic behind the accountability regime is to provide schools with a clear system for success while taking the ability to define success away from schools through standardized testing. This regulatory paradigm only works if there are reliable means for using the system to achieve the

prescribed outcomes. As I find in my observations of Achievement College Prep, there is a gap here that no amount of coupling can fill. They can implement the standards and prepare students for the test, but there is no predictably effective way of teaching the standards. If we want an accountability regime that improves schools, we need to offer instructional methods proven to work for the standards given.

EMERGENCE OF THE SCHOOL ACCOUNTABILITY REGIME

In the wake of *A Nation at Risk*, state and federal governments tightened their oversight of schools by establishing standards for what students should know, requiring a core curricula of science, math, language, and history for graduation; implementing standardized testing; and punishing and rewarding teachers, students, and schools for test scores. These policies constitute what we call school accountability. Common Core was its most recent instantiation.

School accountability is an example of a broader type of change in the way we see the role of government in a neoliberal society, called the New Public Management (NPM). NPM sees the role of government as maximizing social impact while minimizing financial costs (Barzelay 1992; Dunleavy et al. 2005; James and Manning 1996). Agencies should be more like corporations than bureaucracies. Part of this shift is privatization, shifting services from government-run agencies to government-regulated corporations. Accountability replaces administration with monitoring, evaluation, and ultimately capital mobility to ensure services are delivered effectively and efficiently. The government provides service providers with clarity in the form of rules and regulations but retains control over whether the provider has successfully met its goals.

Accountability can successfully lead to improved organizational performance in two ways. Either government is able to move capital to more effective schools or schools become more effective by embracing the regulatory framework. The former has been addressed elsewhere (Arsen and Ni 2012; Lubienski 2003; Zimmer and Buddin 2009). School accountability has largely been set up with the latter in mind. States created standards which define the things students should know. Schools are expected to meet those standards or get better at meeting them.

How standards should connect to test scores has evolved over time (see Hamilton, Stetcher, and Yuan (2008) for an overview). States have been experimenting with standardized testing since the 1960s but policymakers and educators only seriously discussed regulating standardized curricula beginning in the late 1980s (Smith, O'Day, and Fuhrman 1992). The leaders of the standards-based reform movement of the 1990s saw connecting standards to tests as part of a larger reengineering of the educational landscape that would align things like teacher training, classroom instruction, and student assessment.

Reformers realized the difficulty of such a project early on (D. K. Cohen 1990). Standards would not change teaching overnight. The connection between a standard and whether and how best to assess a students' mastery of it was unclear. Moreover, states could legislate standards, but getting teacher training programs to prepare new teachers for the standards and rewarding schools based on test scores were all controversial.

Initial accountability systems themselves were inconsistent (Carnoy, Elmore, and Siskin 2003; Diamond 2007). Some states had their own standardized tests while others used tests from other states. The standards and curricula teachers were supposed to implement were sometimes defined by states and sometimes by local districts. There were often inconsistencies between

curricula, standards, and tests. Moreover, standards and curricula prescribed content but not instruction or pedagogy. Students had to learn about the Civil War and the area of a circle, but teachers could teach them however they wanted.

With this internal inconsistency, it should come as no surprise that, when he looked at how accountability was affecting one elementary school in 1999, Tim Hallett found a battle brewing between a new principal seeking to implement the district's standards and teachers defending their long-standing autonomy over the classroom (Hallett 2010). While the district already had standards in place, teachers had been left to follow them as they saw fit. The new principal changed this by observing classrooms herself, implementing universal grade reporting, and standardizing lesson plan reporting. All of this was too abrupt for teachers who revolted, sending a one-hundred-page complaint against the principal to the district's central office.

Successive waves of policymaking have largely tightened the connection between the various aspects of accountability and increased the punishment for schools that failed to perform. Rather than just send failing students to summer school and putting schools on probation; students would be allowed to move to better schools and, eventually, failing schools would be shuttered altogether.

The accountability system has become increasingly rationalized since it was conceived the late 1980s. It is easier for teachers and schools today to understand what is being asked of them and the stakes for success. That is to say, it is easier for schools to become coupled to the accountability regime (Meyer and Rowan 2006).

TWO DIMENSIONS OF COUPLING

Among education and organizational scholars, schools are famous for their impermeability to regulation and the latest trends in educational policy (Payne 2008). They have been the canonical case of "loosely-coupled" organizations - ones in which what happens in one part of the organization has little influence over other parts (Weick 1976). School accountability thus represents a historically significant shift in the way schools operate.

Coupling occurs along two dimensions: practical and ideological (see Figure 4.1). Most studies of regulation focus on the practical dimension: the extent to which an organization behaves in a way consistent with the regulation. Do they actually follow the regulations? The second dimension, ideological coupling, refers to the extent to which members of an organization adopt the worldview underlying the regulation. Do people understand or believe in the regulation?

Ideological and practical coupling are independent dimensions. Organizations can adopt the worldview of a regime without its practices. For example, in their seminal article Meyer & Rowan (1977) argue that organizations are only coupled to policy symbolically. Leaders adopt new structures or programs as a show of good faith to regulators, collaborators, and customers. But these are only myths about how the organization operates. The daily work of employees remains tethered to the technical requirements of the job. There is strong ideological coupling by organizational elites, but everyone else goes about their business. In the extreme case, you can have what Mehta (2013) calls "paradigmatic" coupling in which organizational members adopt a worldview without any specific practices having yet been defined.

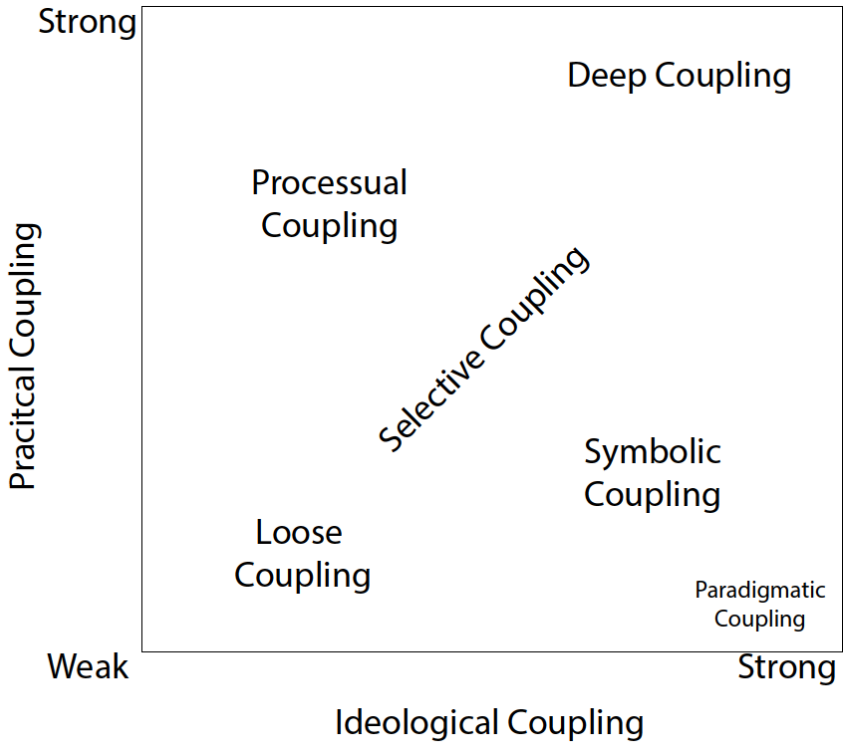

Figure 4.1 Types of Coupling

Organizations can also adopt the practices of a regime without also adopting its worldview (Snellman 2012). In an extreme case, carrying out the practices leads to a regime. For example, Kelly and Dobbin (1998) find that affirmative action compliance offices, created in American corporations during the 1970s, galvanized the creation of diversity management in the 1980s after federal enforcement of workplace discrimination waned. In the absence of a worldview legitimizing their role, these offices continued their work until they could put one together.

These cases of decoupling appear to be rare (Bromley, Hwang, and Powell 2012; Coburn 2004). More often than not, coupling is selective, partially occurring on both dimensions (Pache and Santos 2013). Bromley and Powell (2012) argue that "audit culture" has forced organizations today to respond to pressures from many constituencies, leading them to pick and choose

different programs, practices, and structures to couple to. Hafner-Burton & Tsutsui (2005) show how early policy commitments were used as leverage by social movements to gradually develop countries' human rights compliance regimes. Haack, Schoeneborn, & Wickert (2012) detail the coevolution of practical and ideological coupling among companies adopting corporate responsibility.

These studies demonstrate that coupling is a process in two senses. One, it unfolds over time. At one moment in time, an organization may be loosely coupled or symbolically coupled. However internal and external forces are regularly pushing or resisting policies, practices, and ideas. Organizations change their coupling over time. Moreover, coupling is a process in the sense that it is the result of concerted social action. People work to convince one another of ideas, formulate practices, and hold one another accountable. It is an achievement. Over the past thirty years, school accountability advocates have achieved much by creating a regime that can be coupled to.

The ideal case for accountability advocates is what I call deep coupling – when organizations adopt both the worldview and practices of a regulation. Deep coupling is symbolic of an organization's willingness to go along with policy, but it is also perpetuates itself. Research shows that when organizations have groups of people who believe in a policy and practice it, the organization as a whole becomes more aligned to that policy (Dobbin, Kim, and Kalev 2011; Lounsbury 2001). Deep coupling often occurs among early adopters who adopt a policy to accomplish a pre-existing goal (Tolbert and Zucker 1983; Westphal, Gulati, and Shortell 1997). For those who adopt a policy later, deep coupling is driven by embeddedness in expert networks who can provide technical support for adopting the practices and can explain the worldview

(Kostova and Roth 2002; Simmons, Dobbin, and Garrett 2006; Weber, Davis, and Lounsbury 2009).

The fragmentation and ambiguity in what school accountability involved early on undermined the potential for deep coupling. The evolution of standard tactics, processes, and expectations has only recently made deep coupling possible. Common Core represents a capstone in this sense – a shared worldview and set of practices to organize the work of educators across the United States.

The Common Core Worldview

In his book, *Street Level Bureaucracy,* Michael Lipsky (2010) argued that government programs work by shaping the ways in which the workers who carry out these programs, whom he called "street-level bureaucrats," understand and implement the programs they oversee. Public agencies and officials have significant latitude in defining the ways in which social programs are carried out, laws enforced, and benefits doled out. Government leaders, administration officials, politicians, and policy activists project a worldview with sets of beliefs, values, and norms about how these programs should work, providing cohesion to street-level bureaucrats' discretion. These worldviews, in tandem with the daily life of the program and organizational policies, constitute institutions these street-level bureaucrats inhabit (Hallett and Ventresca 2006).

In education, several worldviews have come in and out of vogue in the past several decades. The most recently studied is the argument put forward by Simon regarding school discipline. Simon (2007) argues that governmental policy has begun addressing social problems through the lens of crime. The central problem for schools is student safety and misbehavior and the way to address it is with crime-fighting - zero tolerance discipline, increased surveillance,

and policing (Hirschfield 2008; Kupchik 2010). This shift in worldview, which Simon calls "governing through crime," led to a variety of new practices and policies carried out by teachers, principals, and police that had real impacts on the lives of students (Nolan 2011; Perry and Morris 2014; Rios 2011).

Common Core represented a return to a worldview in education focused on learning. The problem to be solved was an education system that didn't prepare students for college and careers. Schools did not promote the right kind of learning and the tests were too lenient. Common Core was meant to be a sea change in the curriculum. First, it replaced the specifically American commitment to a wide but shallow curriculum with one that is narrow, but deep common to high-performing European and Asian countries. Students would repeatedly engage a smaller number of topics throughout their education. This repetition would lead to conceptual understanding rather than proceduralization and memorization. In English courses, students would learn to read a variety of texts and pull out relevant information rather than relying on background knowledge or simply spotting techniques like figurative language. In Math, students would learn to identify the types of concepts involved in a specific problem or real-world situation and use a variety of procedures to work towards a solution.

Second, the new standards required new tests. It was not enough to say whether students selected the right answer. The new tests would require them to prove they used the right concepts and procedures to arrive at the answer. Thus, in the Common Core worldview, the new standardized tests were to be a more reliable and rigorous measure of student learning than prior tests. As such, they would be better tools for holding schools accountable.

Common Core Practices

Common Core, in practice, means three things – adopting standards, taking standardized testing, and changing instruction. The ideology of the new standards was made real in the form of new standards in Math and English from kindergarten through the end of secondary school. As before, the standards set out what students should know in each grade level. Compared to prior standards, the Common Core defined fewer standards per grade and largely repeated the same standards from one year to the next. Teachers would teach fewer topics per year and revisit them repeatedly during the course.

For example, the English standards cover four domains: reading, writing, speaking/listening, and language (i.e. grammar and vocabulary). These four domains are repeated every year through grade-appropriate "anchor standards."[1] For example, one anchor standard is "Demonstrate command of the conventions of standard English capitalization, punctuation, and spelling when writing." In kindergarten, one grade-appropriate standard is "Capitalize dates and names of people" and in grades 11 and 12, one standard is "Observe hyphenation conventions."[2] In the minds of their creators these standards represented a sea change in educational curricula, emphasizing depth over breadth; conceptual understanding over memorization.

Two consortia of states were created to develop new Common Core aligned standardized tests. One of these was the Partnership for Assessment of Readiness for College and Careers

[1] The streamlining in math is less obvious as there are eleven domains including "Statistics and Probability," "Geometry," and "The Number System" which are introduced in different ways in different grades. Numerically, there are fewer standards in Math courses at Achievement College Prep under Common Core as compared with the classes under the previous state standards. And Math teachers said that the Math standards represented a narrower range of topic.

[2] See CoreStandards.org for the most current version of the official Common Core standards.

(PARCC). Originally, PARCC planned to develop four standardized tests per year. Two would be formative assessments at the beginning and middle of the year to diagnose students' strengths and weaknesses and help teachers and schools target their resources. Two would be summative, occurring at the middle and end of the year to measure students' learning and hold students, teachers, and schools accountable. States were given the choice of which tests to administer and when. During my focal year, students were given the two summative tests.

Finally, to change instruction, states planned to create resource pools, teacher training programs, and updated certification programs to help teachers teach to the new standards. It is critical to note that, central to the teaching profession in the U.S. is the lack of regulation over instruction. Every profession is founded on the control over some expertise and each seeks to preserve that expertise vis-à-vis other professions and the state (Abbott 1988a). For teachers in the United States, instruction is this sacred ground of expert control – it is up to teachers to decide what the right in-class activity is for their class. The state can require that students learn hyphenation norms, but teachers determine how they learn it. In order for the state to encourage teachers to teach the new standards, they offer instructional material teachers can use, provide educational briefings on the Common Core, work through professional institutions to change their requirements, and rely on the educational marketplace to offer Common Core aligned material.

The Accountable School

In the ideal world of school accountability, the practices and worldview provide schools with a clarity of purpose and direction. Schools perform the practices and integrate the worldview. We should see teachers teaching students the same concepts over and over again

across grades until students grasp them at an abstract level and can use them across texts, contexts, questions, and problems. The ideal school would take testing seriously as diagnostic of their students' learning and a measure of their own effectiveness. During my focal year, I saw all of these things at Achievement. And yet, students performed poorly.

COMMON CORE'S EARLY DAYS AT ACHIEVEMENT

The adoption of Common Core State Standards was part of a broad set of initiatives launched by the Obama administration. As part of the post-financial crisis recovery, Obama initiated the Race to the Top grant program which gave states large grants to improve their education systems. These grants were awarded based on how well states' policies fit with an array of best-practices including encouraging charter schools, evaluating teachers and schools based on performance, implementing data tracking systems, and adopting Common Core.

While few states ultimately received awards, this spurred substantial focus on educational reform at the state level. One senior administrator at Achievement remembers it this way,

> There was just a lot of confusion about what was going on, "is it real?" "Are we doing this?" "How is this really affecting the curriculum?" First, our message for a little while was just, "keep doing the same thing…no one knows what's going on." And I did not get any training and I really should have, but I shy away from PowerPoint trainings that usually just don't do anything for you. Hilary and I did attend one presentation in [nearby city] and I remember us being like, 'this is still a waste of time' It was for Race to the Top which is related to…I mean the whole thing is related but it had like…Race to the Top also had this AP initiative that we were really interested in. But nothing came out of it. So, I would say that very little was done. The state said they were going to use PARCC, more people spent more time looking at what the Common Core is.

The flurry of initiatives drowned one another out early on. Some state-level initiatives survived the Race to the Top process while others did not, making the seriousness of any one of

them suspect. Moreover, the state's standardized test was still based on the previous standards making it difficult for schools and teachers to assess what was supposed to be taught.

In this initial state of uncertainty, the Math and English department chairs decided to do vastly different things. The English department went all in on the new standards while the math department held off.

Archival data from teachers' curricula bear this history out. The following analysis was performed on the workbooks teachers created each year for each of their classes. Nearly all classes had a workbook going back to at least 2009. And, most all of these workbooks contain a template schedule where teachers fill in the daily lesson objective(s) for that day and the state standard(s) to be met. The standards are usually copied verbatum from the state's handbook and are thus traceable across years. I used exact matching to compare standards across each class each year.

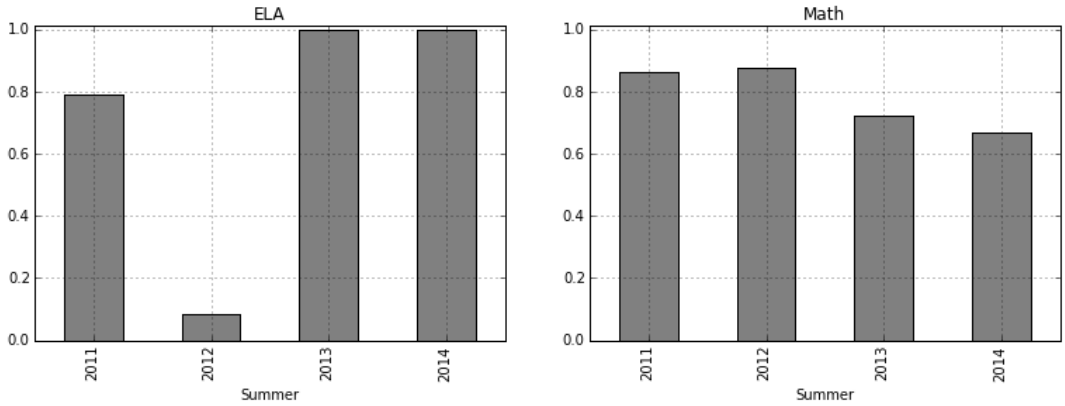

Figure 4.2: Retention Rates of Standards between Years in English and Math

In a normal year, the percentage of standards retained in a course for Math and English was very high. But, only 10% of standards were retained between the end of the 2011-2012 school year and the start of the 2012-2013 school year in English, the year after Common Core was passed. In the subsequent two years, 98% of standards were retained year over year.

In contrast, standards change was more gradual in math. As one senior administrator explained it:

> In the beginning, I think our math results on [state test] had always been stronger than our English results and I think our math department rested on their laurels a little bit. That, because we're doing well on this, it doesn't matter how the standards change or the outside assessment changes, we'll continue to do well. Wherein the [English] department, they were actually excited about Common Core. They were saying, these are the standards we want to align to. We've kinda been waiting for this. There is just more excitement around that piece of it in the English department, where this feels real and authentic and it's not to say there wasn't that in the math department. There was a 'what we're currently doing will still get us to where we are.'

Common Core adoption at Achievement was driven by the chairs of the Math and English departments. In English, the chair had all teachers transition to the Common Core standards as soon as they were adopted by the state. The math department chair left it to individual teachers to address Common Core standards on their own.

COUPLING TO THE CORE

Coupling is not a state one has but a state one accomplishes (Bartley and Egels-Zandén 2016). Through a sustained investment in time and resources, the school gradually became deeply coupled to the Common Core. Teachers gradually adopted the standards and developed a deep understanding of how to teach them in class. The math department hired a new chair whose primary directive was to lead Common Core alignment. The school was an early adopter of the new standardized test and provided external help for teachers teaching to the new standards. By the time I arrived during my focal year, they were already deeply coupled to the Common Core.

Coupling to the Standards in Practice

At Achievement, adopting Common Core standards in practice meant two things – integrating standards into their curricula and developing a deep understanding of each standard. At Achievement, the way you know whether teachers are using the right standards is whether they are written into the curriculum workbook. It is the act of writing the standards down that begins to process of making them practical.

The school practices backwards planning, meaning that teachers create their final and midterm exams first at the beginning of the year based on the standards they need to meet. Teachers then work backwards from there to design their unit and daily lesson plans. Achievement has a standardized lesson plan template teachers follow. As part of this template, each lesson plan has a standard the lesson is meant to teach and a daily objective that accomplishes the standard. For example, if the standard is "students should understand capitalization rules," the objective for the day might be "Students understand how to capitalize proper names." Putting standards into curricula meant not merely writing them down, but that teachers organized everything from daily courses to final exams around the new standards.

While teachers in the English department put the new Common Core standards in their workbooks over the summer of 2012, in math, teachers adopted them piecemeal. Qualitative review of these curricula workbooks confirmed that the entire English department began using the Common Core standards in their workbooks in the 2012-13 school year. And these standards were copied and pasted into the new workbook in subsequent years. In contrast, math department workbooks show some teachers began writing some of the standards in their workbooks in 2012. In the next year, those teachers added more of the new standards and new teachers began adding

the standards for the first time. By the 2014-15 school year, every teacher in the math department was referencing all Common Core standards in their curriculum workbooks.

Standards are not obvious. To teachers at Achievement, they had to be understood deeply. Teachers developed this deep understanding through sustained investigations into the meaning of standards they referred to as "deep dives." In a deep dive, teachers gathered lesson plans, activities, and other curricular material that aligned with the standard and poured over them to better understand what the standard really meant. Teachers often reported doing these deep dives over the summer. As one teacher said, "You gather as much material as you can and do a standards deep dive, but there's really no time for that outside of the summer…I don't have time to sit down and spend seven hours looking at one standard." The school gave teachers summer funding for these deep dives. In fact, deep dives are so important to the school that, when the school faced a budget shortfall at the end of 2014, the only summer funds available went to teachers using the summer to work on integrating Common Core.

These individual deep dives were complemented by focused sessions facilitated by department chairs. In English for example, Ellie, the chair, used department meetings to facilitate a series of deep dives on close reading, the first anchor standard in the new reading standards. The result of these special sessions was a template for planning, implementing, and evaluating a close reading lesson plan to guide teachers in teaching close reading.

The school coupled to Common Core unevenly initially but had fully adopted the standards by my focal year. But adopting standards and integrating them into lesson plans documents were not enough. They had to understand the standards on a deep level. This cognitive understanding is not the same as ideological adoption. The deep dives were about aiding practice. It was about turning regulation into craft.

Coupling to the Worldview of the Standards

Ideological adoption is about understanding the worldview within which the standards make sense and are legitimate. Common Core standards conveyed four beliefs to teachers at Achievement: 1) that the standards emphasized skill development and abstract understanding over procedural knowledge and memorization and 2) a deep understanding of a few topics rather than a shallow understanding of many topics and that this 3) could be done across departments and 4) students progressed through the standards each year.

Compared to the state's former standards, Common Core was portrayed by teachers as emphasizing skills and conceptual or abstract understanding rather than proceduralization, memorization, and broad knowledge. As one math teacher put it, "Common Core is like 'can you cook?' Whereas everything on the state test is 'can you do these three steps?'" One English teacher said, whereas the old state standards asked students to identify the different forms of figurative language like metaphors and similes, Common Core asked students to assess the intent of a text's author. Instead of identifying the tools of the trade, students should be able to infer why such tools were deployed.

Teachers interpreted the decrease in topics not as there being less to teach, but that they could give students more time for each standard. Students would get more "at bats" as they would say. Conceptually, this meant to teachers that the point of the standards was for students to "go deep" or "dig deep" on individual concepts rather than survey a vast body of knowledge. As Matt, the math department chair told me,

> [When I came in for my interview,] David and Kim asked me a bunch of questions [about Common Core] and I was like, "I'm new to this but this seems like good practice, you know, getting students to understand rather than just see." Getting students to dig deep on something rather than learning about something over a small

> period of time. It made sense to me. It's all kind of like, how I had taught and how I had molded my own teaching techniques towards that. So to me, it was very easy for me to buy in. I thought, at first, "Hey we're moving to the Common Core, this is going to be a big undertaking." And it really wasn't.

By shifting standards from domain knowledge to deep, general skills, Common Core also shifted the understanding of the pedagogical goals of different grade levels and departments, providing common ground around a set of basic skills. Instead of seeing courses as just distinct topics – world history, U.S. history, and biology – staff at Achievement saw in Common Core a continuity. All courses taught students to read from certain kinds of texts, write progressively better texts, and approach problems conceptually rather than procedurally. Thus, they picked up on Common Core's ideology of cross-disciplinary and cross-grade integration.

Teachers understood Common Core to have some standards that could unify the school's departments. The administration led school wide trainings on shareable literacy standards like having students gather evidence from texts. This school-wide awareness was complemented by department-level initiatives to teach Common Core-prescribed skills. For example, during my focal year, Ellie, the head of the English department, was also made head of the History and Social Science department and facilitated the same series of close reading workshops with those teachers. The fundamental worldview teachers picked up is that there are a set of skills that transcend fields. As one English teacher said,

> I like [Common Core] because, as much as in my dream world, students walk away from school loving literature, I recognize that that's not always going to happen … But every single student has to be able to read critically, write clearly. And the Common Core is literally just breaking down how to help students do that. So, I like it because these are the skills that people need no matter what field they go into to be successful people. Scientists have to write. Mathematicians have to break things down and understand things.

Moreover, teachers began to see more connections between what students were doing in different grades. Common Core explicitly connects standards across grade levels. In many cases, the same skills and topics are presented with the same standard number with only slight differences across grades (though this is less palpable in math where subfields like geometry and algebra have starkly different topics). This integration projected the image of a unified intellectual progression wherein what students learned in early grades supported their learning in later grades. At the end of the year, all of the staff discussed ways to improve students' transitions from one grade to the next. Common Core made this progression a central tenant of how departments should approach their curriculum.

In an organization decoupled from regulatory ideology, we expect teachers to resist, demean, or simply express a distaste for the Common Core's worldview. Instead, teachers and staff at Achievement adopted the worldview being projected by the Common Core standards and felt they were just "common sense." They were not fanatics about it. Strong ideological coupling does not entail rabid, doctrinal faith. Instead, it can include the sophisticated, appreciative, and skeptical endorsement embodied by teachers at Achievement. They did not deny the importance of procedural knowledge and the usefulness of being broadly knowledgeable. Nor did they redesign every class to teach Common Core standards. Teachers and staff understood there were trade-offs being made and believed Common Core represented a better balance. For them, the world of deep, skills-based learning as the basis for teaching across grades and departments was the ultimate goal toward which the standards were guiding them.

Teaching for the Test

The perennial critique of standardized testing is that it encourages schools to teach to the test. For example, students are taught strategy for answering multiple choice questions rather than the causes of the American revolution. This represents a shallow view of the ways in which Achievement prepared its students for Common Core testing. Achievement was most concerned as an organization about closing the achievement gap in test scores. Thus, passing tests was the school's primary objective. And testing itself was a central part of the life of the school.

Achievement was putting their students through standardized testing on Common Core standards before PARCC released its first test. Achievement was working with a testing company called Achievement Network or ANet for short. ANet tested students at Achievement four times per year assessing their mastery of the state's standards and providing teachers and administrators with student's scores on each standard. The purpose of these tests was to provide no-stakes feedback on how well students met the standards and were used to adjust course content. When the state adopted Common Core standards, ANet adapted its tests to evaluate students according to the new standards. As the PARCC began releasing practice test items, ANet began adjusting its tests.

PARCC's first practice tests were also a boon to teachers at Achievement. As one teacher said, "You think you understand [a standard] and you see something else or you see how it's going to be assessed on PARCC and you're like, 'Oh, that's really what they meant.'" Teachers in math and English were given stipends to spend the summer of 2014 reviewing PARCC practice tests and interpreting what the assessments meant for their classes. In their class's archives, four teachers document this work in the form of summaries reported to the school's leadership. In these reports, teachers outline the meaning of the changes made by the Common

Core standards, enumerate the skills and question types emphasized in the practice tests, and assess how to make their classes more Common Core aligned.

When the state announced it would begin giving the PARCC exam as an option to schools during the 2014-2015 school year, the school's administrators were quick to act. The state asked schools whether they preferred to take the PARCC or the traditional state test for the grades and subjects where both tests were offered. The school's leadership elected to participate in the PARCC test. As one senior administrator said, by being "early adopters" of PARCC, "we could practice that earlier."

The first tests were given in late March. As they drew near, teachers began their test prep. Figure 4.3 shows a side-by-side comparison of mentions of "PARCC" and "Common Core" in department meeting documents in 2015. While Common Core was mentioned for most of the year, PARCC was a sprint from February to April during which time the school did the traditional teaching-to-the-test.

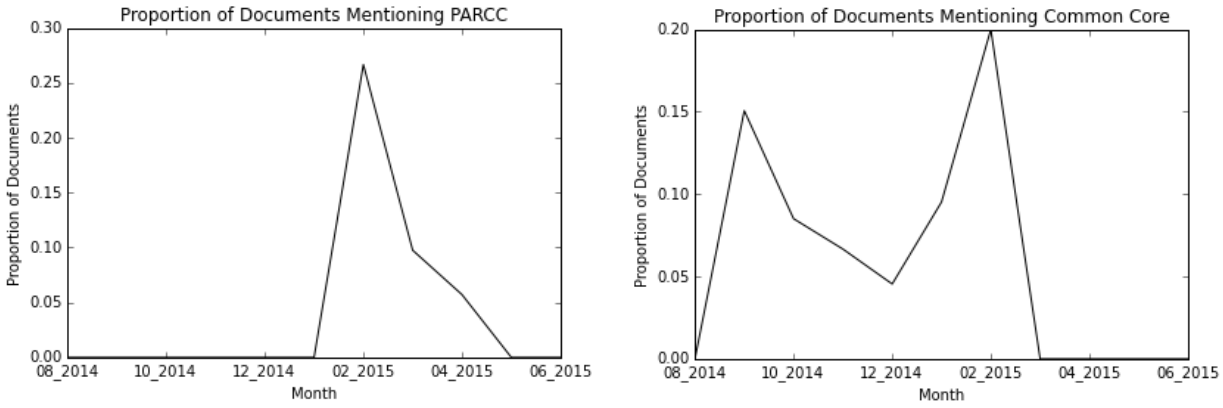

Figure 4.3: Plot of Mentions of Common Core and PARCC Testing in 2014-15

In the run-up to the tests, teachers focused on reviewing the standards, preparing students for PARCC-style test questions, and getting students intellectually and emotionally prepared for standardized testing. Teachers, particularly in math, adjusted their curricula to cover missing standards or review standards on which students had performed poorly during the year. Teachers also incorporated PARCC-style practice questions into their daily lessons. For example, Gloria, a middle school math teacher, gave her students questions from the PARCC practice test as a warmup activity at the beginning of class.

Finally, teachers in English devoted department meeting time to putting together a skit about test taking strategies. While I never saw 2015 video, the video from the prior year featured teachers slaloming down a hill, staged on the school's stair cases. (The 2014 Winter Olympics had just finished.) Each gate represented a step in their test-taking strategy: "prepare," "read," "evidence," "respond." The first teacher missed a gate and fell, disqualifying her from the race while a second teacher passed through all four gates, winning the gold.

Faculty and staff at Achievement College Prep were committed to the PARCC test. They tested their students against Common Core standards early, adapted their teaching materials to fit the new test, were early adopters of the PARCC test, and spent a substantial amount of time preparing students to take the test.

PARCC was supposed to be a more rigorous test that better reflected students' readiness for college. This resonated with teachers and staff at Achievement who saw their students excel on state tests only to then struggle on college entrance exams like the SAT, ACT, and Advanced Placement tests. They believed that the state's standards were too low and that a more rigorous test would show the tough, but necessary reality. As one teacher said,

> [The state test] is not rigorous. We know that. So the fact that 100% of students achieve proficient on [the state test] isn't that impressive. It's the first time we're seeing results on a test that has the standards we want to be reaching. And they're just not reaching them. Those are the standards students need to be reaching to get into these better schools, right?

It is important to note, to teachers the standards did not determine the amount or quality of education students were expected to be achieve. The tests did. As one teacher said,

> I think that it's great that the bar is higher across the board. … I think it's great. It's a good reality check. I think it's good for everyone to have a national standard. And I think that's what PARCC [provides], because Common Core standards aren't telling you how you're being assessed and the level of rigor. So you need something like PARCC or another option to say, 'this is actually the bar for what this looks like.'

Standards are not transparent tools. They need tests to give them a concrete, meaningful reality. Moreover you can teach whatever you want and believe students understand it, but, according to staff at Achievement, the standardized test will tell whether you are, in fact, right.

Ironically, the PARCC itself was not modelled on any of the existing college entrance exams. It was built from scratch. In fact, assessment design was flowing in the opposite direction. The theories being used to create PARCC were beginning to be adopted by the college entrance exams. The SAT, ACT, and AP were themselves planning to become more like the PARCC. Teachers and staff knew this. Even the state's science curriculum was being re-written to be more like Common Core. Thus, the idea that PARCC would diagnose students' college readiness was not based on the PARCC mimicking other collegiate tests, but based on the belief that it was a more challenging test measuring deep understanding and skills, which educators considered a better college preparation than broad knowledge and proceduralization.

Outcome

You can hardly expect an organization to try harder to succeed at being regulated than Achievement College Prep. Though the math department waffled early, teachers and administrators eventually picked up the slack. During my focal year and the first year of PARCC testing, they were ideologically and practically committed to teaching their students the Common Core standards and preparing them to pass the newly designed PARCC test.

This did not seem to do them much good. Their scores were disappointing. Fewer than half of the school's students received a passing grade of 4 or higher on any PARCC test. On most tests, the percentage of students passing was in the twenties and thirties. Their test scores dropped from the top 5th percentile to the 50th percentile for schools in the state. But they did not challenge the test or the standards. They owned them as part of the ideology of the more rigorous and accurate exams. The results fit the established idea that their students were unprepared for college. They committed to working harder. As is likely only found among deep coupling organizations, their failure served only to deepen the coupling process.

The reason they performed so poorly, as they will tell you, is that incorporating the standards and preparing for the test are not enough to ensure good test performance. Students must be taught well. However, it is not clear to anyone what it means to teach Common Core standards well.

At the end of my focal year, I asked each interviewee how they thought Common Core implementation had gone. Some answered before the test results were in, a few afterward. Every person gave me a different answer. Some saw implementation as a short-term adjustment to the new law. For example, one math teacher said, "I think in math and English, we're probably at

like 80-97% of implementation. I think every teacher in every subject area is very aware of the Common Core standards and knows what they should be doing."

Some saw implementation as requiring a shift in instruction. It wasn't just teaching new things, but teaching them in new ways. One English teacher said,

> I think it's been okay. Implementing Common Core in our writing instruction has been pretty good, given that that was our focus in the past few years. So that makes sense that that would be better. I think implementing it in our reading instruction is still shaky. And I think it's really about the teacher's investment in really understanding how to implement specific parts of the language of the standard and not something that we've really focused on as a department.

Finally, others saw the new regime as requiring a shift for everyone across the school. Teachers not only needed to teach students to understand the concepts behind linear equations and triangles, but all teachers needed to teach students that deep understanding was the goal of all classroom activity from sixth grade to graduation. One math teacher said,

> Staff – Yeah, it's a bit too early to tell. I think it's a seven-year process.
>
> Interviewer – And why seven years?
>
> Staff – Because it's seven grades. All of a sudden…these kids have been, for years, not asked to do the majority of the thinking, now I'm asking them and they're like, "Just tell me what to do." And I'm like, "No, I want you to discover it" and that was a big thing. This year's seventh graders, after having it for a year, are going to be more used to it. So, by the time they get to upper level classes and are asked to look at a mathematical concept and develop it on their own, they're going to have A) the math background and B) the experience of looking and [saying] "Hey, what am I doing? I'm stuck."

Each of these points of view point to vastly different ways for succeeding in the state's accountability regime. In the short-term view, standards adoptions and test prep is enough to reach Common Core compliance. In the mid-range view, teachers' instruction needs to be more Common Core aligned through deep understanding. In the long-range view espoused by the math

teacher, the standards require an entirely new pedagogy before students are able to learn the way implied by Common Core.

The variety of responses to "how far do we have to go" is illustrative of a basic problem with Common Core and standards-based accountability: they provide no guidance on how students should be taught. This lack of guidance has been a cornerstone of accountability policy. Legislators agreed not to legislate how teachers teach. However, different standards and testing protocol implicitly favor certain forms of instruction. According to school administrators, the school's "drill and kill" pedagogy, done largely through high volumes of quiet, independent work through packets of pre-printed problems, would not prepare students for the open-ended writing, argumentation, and abstract understanding required by Common Core. But it was unclear to them what the alternative should be. This lack of guidance guarantees failure even among those most deeply committed to the standards and test.

The Search for Common Core Alignment

The deep dives and test preparation proved to staff the need for new instructional material early on. However, because Common Core was new, no one knew what instruction they should be using. They developed their own material as departments, but they also brought in external resources like textbooks, lesson plans, coaching, and consultants. These resources didn't referee what was right or wrong, effective or ineffective, but what fit the "sprit" of Common Core, in their words, what was Common Core "aligned."

Teachers themselves spent a lot of time working on new lesson plans, activities, and techniques they believed would help students master the new standards. As I mentioned above, the English department spent several months working on a framework for teaching close reading

which they would use to evaluate lesson plans and activities. In math, Matt spent a lot of his time with teachers helping them think through how to teach a standard from different angles. This "figure it out for yourself" approach is essential to teaching as a form of craft. Good instruction comes from the autonomous work of experts. But it takes a substantial amount of time. And there's no guarantee that a teacher will get it right.

Teachers and chairs looked to trusted sources outside the school for shortcuts – activity ideas, lesson plan templates, and whole curricula written for Common Core that could provide the right instruction. One type of source for Common Core aligned material was the open access clearinghouse. The one most often cited was EngageNY – a free clearinghouse of material created by the State of New York in 2011 to help its schools succeed under the new Common Core standards. A second source, particularly in math, were textbooks. Matt's first goal for the 2015 school year was to have every teacher using Common Core aligned textbooks.

A final external source of materials for teachers and chairs was visiting other schools. The school coordinated annual visits to other schools in the area so their teachers could see how other teachers and schools worked. Moreover, working within their charter network, the school organized special visits to high performing schools in other regions for a select number of teachers and leaders to bring back effective methods. Teachers did not adopt curricula and activities unthinkingly. Instead, they cherry picked what they thought would work or went to them when they needed inspiration or guidance for a particular standard. In essence, they used their expertise to make judgements about what they thought would work for their classes.

The school also brought in new staff, coaches, and consultants. In my time at the school, they hired two new department chairs based on their ability to support teachers' adjustment to Common Core instruction. This was particularly true in the History and Social Science

department where the new chair was hired their expertise in teaching literacy and had background in either history or social science and had never taught children.

Teachers, chairs, and administrators worked together to hire coaches with outside expertise in Common Core aligned teaching. They brought in coaches from ANet as well as a local nonprofit to coach teachers in writing instruction. Chairs and teachers were the ones who decided to hire coaches while administrators found the money to pay them. Administrators on the other hand took the lead in hiring consultants.

Consultants were the way in which the administration exerted direct control over Common Core adoption. In principle, administrators gave department chairs control over teachers' instruction. Departments were the subject matter experts on how to teach their subject, not administrators. Consultants were the administrators check on this power. As one administrator said,

> Listening to the math experts in the building at the time, there was this sense that what we're doing now is okay but [in] getting smarter on that side of the house and talking to other people and looking at their curriculum and looking at what was going on around [the city], having these outside individuals come in and do audits of our curriculum and have them say, it's not close to where it needs to be, lit the fire under…[trails off]. And it just happened that it was at a transition point that really gave a clean slate to say, 'we really need a new department chair and we need that person to charge forward on this.'

This administrator's own knowledge was not enough to challenge the math department chair. They needed to have other experts who could say that the curriculum wasn't right.

I was privy to one visit by consultants brought in to audit the math and English departments near the end of the year, between the two waves of PARCC testing. The consultants popped in and out of classes throughout the morning to evaluate teachers' instruction. After

lunch, they met with the department chairs and school administrators to debrief. In that meeting, the lead consultant recapitulated their mission this way,

> We talked with David [the Dean of Instruction]. What he asked us to do is to go into the classroom thinking about Common Core especially with your [points to Ellie] priorities around questioning. Math: increasing rigor and breaking from processes. We start with EngageNY's work where they developed a set of good actions you should see around Common Core. So, we were only in [each class] for 20 minutes, so this reflects those actions that are easy to see in 15-20 minutes and what's challenging. So, we can't give you feedback on everything.

The administrators set the agenda for the audit by charging the consultants with in-person evaluation of teachers' in-class instruction. The chairs shaped that agenda by defining which aspects of instruction they wanted to ensure they got feedback on.

The substance of the audit itself revolved around what one should see if a teacher's instruction is Common Core aligned. I've been calling this the ideology of Common Core – the types of things we expect to see because of Common Core, not what has been explicitly legislated.

> Bailey, the consultant for math, began, "We focused on three areas 1) that the work of the lesson reflects shifts in the standards. 2) Teachers practices. And, 3) all students have the opportunity to exhibit the math practices. Our headlines, students need to be engaged in a few, rich tasks. And teachers can move from the procedural into the conceptual. Some examples are, in the classroom, the assessments, the work in the packets is really scaffolded. Where there is real-world application, but the focus the students have is still procedural." "Which grade?" Matt asked. "It was an overall theme," Bailey answered, "but especially 7 and 8, but some in 6 as well. Do the teachers know what the standards look like in their teaching? Do they know what the core concept is underneath the core focus? If they do, then the structure of the curriculum is holding them back. So, it's the knowledge of the teacher or the structure of the curriculum." Matt responds matter-of-factly, "The teachers' knowledge is fine, so it must be the way it's set up."

The consultants' feedback maps directly onto the Common Core worldview. Instruction should be deep, focused on "a few, rich tasks;" conceptual rather than just procedural, and more

rigorous and less scaffolded (a term of art for making concepts easier to understand). The consultants' role was not to fix these issues. They did not leave behind curricula or lesson plans. Instead, they were the referees for whether or not instruction was Common Core aligned.

Achievement College Prep largely sought to make its instruction fit with Common Core by going outside the school. Teachers and department chairs brought in material from textbooks and digital repositories. They hired coaches, department chairs, and consultants who could help teachers develop their classes. Few of these organizations, repositories, or instructional activities were new. They existed before Common Core and were adapting themselves for the new regime. But none of them were provably effective for Common Core. The PARCC, which could determine whether some instructional strategy or curriculum was effective, had only just been invented. Instead, they were evaluated according to whether it they were Common Core aligned, that is, whether they fit the goals of Common Core.

DISCUSSION

Education's Marshmallow Challenge

The Marshmallow Challenge is a team-building exercise made famous in a TED talk by Tom Wujec (2010). Teams have 18 minutes to build the tallest tower they can using only a piece of string, some tape, and a handful of uncooked spaghetti noodles. The challenge is that the tower must support a marshmallow at its peak.

The task seems feasible, if unusual. It's clear the brittle noodles won't sustain much weight and tape and string won't create the strongest joinery. The fulcrum of competition too is sensible – height is clearly difficult to achieve with the heavy marshmallow at the top.

According to Wujec, a third of teams fail to create a structure that can stand. No one knows how to build marshmallow towers out of spaghetti. Adding payment for winning actually stifles experimentation, prototyping, and testing. Participants approach the challenge boldly. It's "go big or go home" and it all falls apart.

This is the problem with Common Core – teachers are given the standards and some course material (the noodles and tape) and expected to produce students who pass standardized tests (the tower with a marshmallow at the top). But there are no instructions about how to build the tower or what kinds of noodle structures will hold together.

The flaw in the accountability system is the gap between regulation and evaluation. The accountability regime simply sets a bar and tells teachers to jump. It leaves the practice of teaching up to teachers which, in reality, means to the marketplace of consultants, textbook makers, and free resources. This market is rife with un-tested programs, packages, and procedures, each promising to bridge the gap between Common Core standards and high test scores, and no way of ascertaining their effectiveness.

Normalizing Failure

One of the founding criticisms of education leveled by the no-excuses movement is that failure is normalized in the traditional American education system. I argue failure is baked into the system. Achievement College Prep's experience shows the game cannot be won. The task we set for schools is impossible given the means we have. But some may disagree.

Some may argue that this was a down year for Achievement. Every year presents its own unique challenges. However, a failure rate of 50-80% is much too large to just be a bad year. With rates this low, it is hard to imagine a good year producing a failure rate in the single digits.

Further, if it is a bad year, it should have been clear to everyone what the problems were. If schools really have a clear means for producing Common Core-educated students, the "bad year" explanation should be apparent. Instead, diagnoses differed.

Some may argue that students at Achievement really did not know the standards. Teachers and staff at Achievement seemed to agree with this. They believed the test was generally accurate in part because it lined up with their experiences with the SAT, ACT, and AP tests. The PARCC did what it was supposed to do: reveal what students really know by offering a more difficult test. Indeed, tests are meant to be diagnostic. However, this is different from tying them to accountability. We test to ensure schools teach students effectively. Punishing students, teachers, and schools for failing tests they are expected to fail is not accountability. It's merely performing social inequality.

Moreover, the state's analysis showed that students' scores on the PARCC were no more correlated with college entrance exams or college grades than students' scores on the prior state test. The notion that the test was more college-aligned has not been borne out by later studies. Moreover, the college entrance tests themselves were undergoing their own re-designs to be more Common Core aligned. In reality, college entrance test scores have historically been more indicative of a student's social background than their potential for success in college (Grodsky, Warren, and Felts 2008).

Both responses, that even deeply coupled schools can have bad years and the test was valid, miss the point. If standards and testing are meant to drive school improvement, then schools must be able to effectively respond to them by improving. The lack of proven instructional methods makes this impossible. The goal of accountability is not perpetual mediocrity, but the policy regime we have in place guarantees it.

CONCLUSION: REGULATING SUCCESS

Under the New Public Management, the state gives over control of its agencies to private actors in exchange for more administrative effectiveness or efficiency. Accountability mechanisms ensure these agencies meet their basic performance goals. This situation sets organizations up for failure if they cannot control their performance. For this form of change to work, following the regulations must lead to success.

In most cases, regulatory compliance leads to success. We have speedometers in our vehicles and, if we stay under the speed limit on our speedometer, we will also stay under the speed limit on a police officer's radar gun. Some regulations are less certain but just as feasible. The Environmental Protection Agency requires automobile manufacturers to achieve a certain fuel efficiency standard in their fleets. Perhaps manufacturers can't guarantee they will achieve especially ambitious goals or be very precise in what reductions are possible in fifty years. But they have reliable mechanisms for cutting fleet emissions which, if automakers decide to couple to the regulation, can effectively cut emissions consistently and persistently.

This is exactly what school accountability cannot do. Without predictable mechanisms for teachers to teach students to the standards, we cannot expect the school accountability regime to lead to consistent and persistent gains in student learning. The No Child Left Behind Act's expectation that schools get 100% of students to pass their state's standardized test in 12 years was fantasy.

CHAPTER 5

CHURN AND BURN AND THE PURSUIT OF EXCELLENT TEACHERS

INTRODUCTION

Organizational change is not always directed by a clear strategy controlled by the organization's leaders through formal policies, procedures, and practices. Sometimes it emerges from many different decisions across time and under various constraints. This helter-skelter form of organizational change has been referred to as organizational anarchy or, the "garbage can model" of organizational decision-making (Cohen, March, and Olsen 1972). Despite the image of an unstructured mass of activity, the result of garbage can decision making is not disorder or necessarily fragmentation. Rather, patterns of organizational action *emerge* from the many local changes. Based on my observations at Achievement College Prep, I argue the employment system, referred to in the popular press as "churn and burn," emerged as a result of an anarchic change process.

"Churn and burn" refers to the system wherein some schools, typically no-excuses charter schools, hire young, less-trained teachers who work longer hour at lower wages, burn out, and quit teaching within five years. No school founder or dean or charter management organization sat down and said, 'this is what we want to do' (hence I call it a system rather than a strategy).

Instead, at Achievement, churn and burn resulted from a variety of different decisions in which school leaders did not have all the answers and were unsure of what their strategy should be. There was no *clarity* in the process of creating and shaping Achievement's employment

strategy. There was also little control. The school's choices were constrained by the job market for teachers and their budget. And the school failed to leverage the things they did have control over like teachers' work requirements and class assignments. Just because an organization can control something, does not mean it exerts that control.

Churn and burn emerged as a relatively coherent system because the school's leaders had an implicit ideal teacher in mind when solving each of these problems. This ideal employee who "fit" the organization is common in the postindustrial workplace where the service sector dominates, and work is emotional, social, and cognitive rather than physical. They hired a certain *kind* of person and established a set of working arrangements to create a workforce that would work for long hours at less pay and ultimately burn out.

CHURN AND BURN AT CHARTER SCHOOLS

Charter schools were created, in part, to improve public education by changing the way teachers were recruited, retained, and compensated. The portrayal of the ideal charter school was to recruit the best and brightest to become teachers, provide intensive on-the-job training for these wunderkind, and reward the effective and fire the ineffective (Furgeson et al. 2012; Hoxby 2002; Neal 2011; Podgursky and Springer 2007). To pursue this ideal, charters were given waivers to recruit teachers from non-traditional training programs and exceptions from collective bargaining agreements with unions which regulated teacher compensation, retention, and work hours (Podgursky and Ballou 2001). In practice, charter schools generally achieved this ideal on average (see Furgeson et al (2012) for an overview). However, this system has also led to a common set of dysfunctions.

In the ensuing thirty years, research consistently shows that the working conditions for teachers are worse at charter schools than traditional public schools. Charter school teachers are paid less than their traditional public school colleagues. Part of this is due to the average tenure of charter teachers (Malloy and Wohlstetter 2003). Young charter teachers make more than their equivalent public-school counterparts (Podgursky 2006). However, charter schools do not pay senior teachers as much as public schools. In his review, Podgursky (2006) found that the most seasoned teachers at public schools were paid roughly one-third more than teachers with the same experience at charter schools. Thus, charters in aggregate pay less than public schools and pay veteran teachers less than their traditional public-school counterparts. Notably absent in this pay comparison has been inclusion of benefits like retirement and health care which are typically generous for teachers in traditional public schools (Olberg and Podgursky 2011).

In addition to lower pay, charter school teachers typically work longer hours (Stoddard and Kuhn 2008). These hours come from longer school days, longer school years, and more work requirements (Hoxby 2002; Malloy and Wohlstetter 2003; Podgursky and Ballou 2001). Hoxby (2002) found that charter school teachers worked an average of 10.9 hours of overtime, or a 51 hour week, compared to the 6.4 hours (a 46 hour week) reported by public school teachers and were assigned more responsibilities than public school teachers.

Charter school teachers also work with substantially less job security than teachers in traditional public schools. Unionized teachers in public schools typically receive tenure after three years of work whereas most teachers in charter schools work on annual contracts or at-will (M. Podgursky and Ballou 2001). Only two-thirds have a salary schedule setting predictable rates of pay over time compared with 88% in traditional schools (Podgursky 2006). Teachers in

charter schools are much more likely to be fired than teachers in traditional public schools. When they are fired, they also lack recourse to contest their disemployment (Stuit and Smith 2012).

These poor working conditions correspond to high rates of burnout and turnover (Torres 2016). Teachers at charter schools are 50% more likely to leave each year than teachers in traditional public schools (Keigher 2010). This gap appears to have shrunk recently but remains (Goldring, Taie, and Riddles 2014). Charter schools have higher rates because of staff composition and school characteristics (R. M. Ingersoll 2004). Stuit and Smith (2012) show that such variables could only explain 60% of the difference. Teacher-level characteristics like tenure, training, and part-time status explained 42% of the gap while personnel policies like retention, pay, and union membership explained 17%. Furthermore, teachers who leave charter schools are more likely to cite dissatisfaction with their working conditions as a reason for leaving (Miron and Applegate 2007; Stuit and Smith 2012).

Charter schools were not meant to churn and burn their teachers. So how did this arise? Evidence across these studies suggest that while the aggregate effect is such that charter schools generally are worse places to teach, the variation between schools is high. Podgursky (2006) found that most of the variance in pay among teachers at charter schools occurred between schools rather than within them. This indicates that most of the variation in pay among charter schools happened as a result of different schools having very different policies. This contrasts with public schools which typically have compressed salary ranges. The quantitative evidence suggests pay policies are largely set by individual charter schools and vary substantially among schools. This fits with Hoxby's (2002, 882) qualitative evidence on charter administrators' response to questions about wage setting. Many administrators adopted district pay scales but gradually adjusted the scale to meet their priorities.

While churn and burn is an aggregate pattern, it is especially prominent in some charter schools and not others. Chief among the exemplars of churn and burn are no-excuses charter schools (Lake et al. 2010, 7, 58; Wilson 2009). Lake and colleagues report that teachers at no-excuses schools average more than 60 hours a week. At KIPP, a national network of no-excuses charter schools, teachers were required to be in the building from 7:30am to 5pm and then had to be on call thereafter to help students with homework. In these schools, turnover can be as high as 35% annually.

Thus, while churn and burn is common in aggregate, there is a great diversity in actual practice across charter schools. To understand how churn and burn has emerged, we must look to the schools themselves.

CHURN AND BURN AT ACHIEVEMENT

Achievement exemplifies the churn and burn system. Most teachers did not study education in college but instead come through alternative paths. They were hired based on their commitment to the no-excuses mission, skill at classroom management, and coachability. Once hired, the workload combined with the tight coupling of the professional community and focus on urgency lend them working 60- to 80-hour weeks. Every teacher is hired on a one-year contract and offered a new annual contract after a formal evaluation, a process referred to as being "asked back." Finally, all teachers have the same responsibilities and receive the same salary regardless of experience or performance. While wages are competitive initially, they quickly become less than public schools.

The market logic for hiring teachers stops at the school door. Achievement's strategy for employing teachers is not built around the philosophy of attracting the best and brightest,

developing top performers, and releasing the low performers. There was no coherent vision at all. Instead, the many pieces of teachers' employment were created and developed independently. What they all shared however was an assumption about the kinds of people who would teach at Achievement. They believed that teachers were 1) "Type A" people willing to sweat the small stuff; 2) Achievers, willing to work to become the best, and 3) True Believers, willing to give up traditional benefits for the greater cause of helping disadvantaged students. These beliefs cohere as an ideal-type charter school teacher. Not all staff see themselves as meeting each part all the time. Instead, it persists in implicit assumptions, formal expectations, and informal reputations as an idealization of teachers and conscious belief about the "kind of people who work here."

The ideal-typical teacher is someone who works long and hard to perfect their craft in pursuit of the school's mission. This image did not arise at the founding of the school and then guide the construction of the school's employment practices. Rather, different pieces of the ideal type were called forth at various times to address specific questions and problems. The impact of this uniform belief on these disparate decisions has been to engender a unified system linking hiring, development, retention and advancement into the churn and burn model.

HIRING HIGH ACHIEVERS

Charter schools exhibit high degrees of churn and burn in part because they hire younger and less experienced teachers who are likely to leave (Ingersoll, Merrill, and Stuckey 2014). However, charter schools exhibit higher rates of churn than just attributable to teacher inexperience (Stuit and Smith 2012). There is more to the story. I argue in this section that Achievement sets themselves up for churn and burn by hiring teachers who fit the ideal type.

If you want a job at Achievement, you must have three things: basic technical competence in classroom management, a commitment to self-improvement and professional growth, and genuine belief in the school's no-excuses mission. These create a population of eager achievers who lack experience but are ready to learn and motivated intrinsically to do what it takes to teach disadvantaged students. This process begins in how they source teachers. Teachers at ACP come from special programs and they rigorously test candidates on the three characteristics.

We see the logic of the ideal type teacher re-enforce itself during the focal year as they try to make a push for more diverse teachers. The goal was unaligned with the vision for their teachers. School leaders were not particularly convinced that a diverse teaching staff would be more effective than their largely young, white, female workforce. Instead, diversity was seen as a way to achieve a more inclusive school community, empathic faculty, and tighter bonds between teachers and students. Rather than choosing diversity over performance (changing goals) or performance over diversity (decoupling), they believed they could do both. In the attempt, we see the logic of the ideal teacher reassert itself for teachers of color.

Sourcing High Achievers

Most teachers hired by Achievement did not start their adult lives interested in teaching. They mostly went to four-year colleges and universities for degrees in political science, engineering, or biology. Many, late in their journey through college or shortly thereafter, felt a calling to take what they had learned, whether the causes of racial injustice, the origins of life, or the love of physics, and share them with students who they perceive are not given the opportunity or are not expected to learn them.

Because teaching requires certification, most people who come to education late are routed through a master's degree. The teachers whose route led them to Achievement enrolled in an alternative certification program started by the founder of a no-excuses school and designed to feed no-excuses schools in the region. I refer to this program pseudonomously as the No-Excuses Teacher Training program or NETT. NETT was an unaccredited education program operated by a charter management organization and offering a master's-level degree in teaching. The selling point for NETT was as a gateway into teaching disadvantaged students in charter schools.

The program was designed to be a technical bootcamp for teachers, focusing on training people in the unique style of no-excuses teaching: high paced, quick to discipline, and focused on classroom management and student compliance. Pedagogy was absent. Management mattered. Would-be charter teachers spent the first six months of training in simulated classrooms, then spent six months as an apprentice teaching and coaching, almost always in a no-excuses charter school. Teachers receive their degree only after spending their second year working as a full-time teacher, again almost always in a no-excuses charter school, and receiving approval from a NETT coach observing them throughout the year.

At the beginning of my focal year, the school hired eight new full-time classroom teachers. Of these, five had just finished the NETT program. The other three were veteran teachers. In an interview, one senior administrator described the school's rationale for hiring so many NETT graduates:

> Last year and the year before, we hired a lot of NETTs with the idea that, the thinking was, they're being trained with a lot of the methods, systems, planning, all of those things. They work them really hard, really long hours and so as a teacher prep program, they're very aligned to Achievement and felt that some of the teachers have and hopefully will be really good fits for Achievement.

NETT is part retraining program and part breaking in process. It ensured that its graduates had the basic skills (and temperament) for sweating the small stuff with urgency, were receptive to personal growth through coaching, and committed to the no-excuses vision of educating the disadvantaged. It also prepared teachers for the lifestyle and workload required to make it in a no-excuses charter school.

The Problem to be Solved - Diversity

The emphasis on teachers trained to work in the no-excuses system was challenged during my focal year when students walked out of school in solidarity with the Black Lives Matter movement. In discussions following the walkout, students and faculty raised the issue of staff diversity pointing to the fact that the school's staff was mostly composed of young, white women. These discussions echoed long-running observations of the homogeneity of the school's staff and many critiques by different students and staff about the lack of men, or people of color, or people with experience. Indeed, when I started observations in 2014, five of the eight new, full-time teachers were young, white women. Two were veteran, white men. The one newly-hired woman of color was a NETT student.

In observations and interviews, staff and students noted that when the school hired people who were not young white women, they tended to be placed in part time or low-status, non-teaching roles as assistants, substitutes, and office support staff. In fact, while most of the full-time teachers who started in 2014 were young, white women; *eight out of nine* non-full-time teachers hired were either people of color (five) or men (four). Only one person hired in a non-full-time teaching role was a young white woman.

In the 2014-2015 hiring cycle, administrators decided to focus on hiring diverse staff. The problem was that the NETT program was largely made up of young, white, women. As one senior administrator put it:

> In terms of staff diversity, and the staff at NETT will be the first to say, there are a lot of white females. And so, the diverse staff from there is not as high typically which is something they're working on as well. And also, they're all first-year teachers. When we talk about diverse staff, we're talking about gender and race, but we're also talking about teacher tenure and how can we hire staff that have more experience coming in if possible.

The school had to turn elsewhere to find diverse staff. They expanded their search, participating in job fairs and joining other schools in the city in a partnership to bring diverse teachers into the city. Expanding beyond NETT represented a substantial challenge for the school. It required the school's time and energy to attend special diversity career fairs. They had to compete against many other schools looking to hire in the relatively small pool of diverse teachers. The teachers at these job fairs came from a broad range of educational trajectories, backgrounds, commitments, and beliefs.

Given the competitiveness in hiring teachers of color and without NETT to filter candidates for fit, it would have been easy for school administrators to simply hire diverse teachers and sacrifice their commitments to hiring teachers based on their beliefs, develop-ability, and focus on classroom management. Instead, as we expect in emergent processes, the idea of what a good teacher is shaped the way they sought to solve the problem at hand.

The Ideal Type in Action

I was able to observe several conversations among administrators about individual candidates. In these conversations, the goals and purposes of increasing diversity never came up.

Administrators recruited diverse applicants but evaluated them based on their core commitments. In fact, the goals of diversity – increasing empathy between students and faculty and building warm relationships – were cited as marks against candidates.

In a meeting of department chairs during the spring hiring season, they discussed whether to promote an assistant teacher of color:

> When asked about promoting one of the assistant teachers to a full-time role, a couple of chairs had questions: Sandy asked, "Where is her trajectory?" Robin followed quickly, "And her philosophy?" "Where is her performance?" Ellie asked. David, the dean of instruction, began to answer each one, "I don't know about the philosophy piece. I don't have my notes from her mid-year [review]. I remember her saying, 'I want to be a teacher. I want to be here.'
>
> Robin, the science chair said, "I had a conversation with her a month ago and she said she didn't believe in a no-excuses charter school." "Well, we should definitely follow up with that," David replied solemnly.
>
> "How was the observation?" Ellie asked. David responded, "Her pacing was way off, but she was smiling more, she was trying to project more enthusiasm." "She did a nice job," the director interjected. David added, "You could see more warmth and more joy and excitement with kids. And she pointed out that pacing was an issue for her. So, she's self-reflective."
>
> "I think some staff would resist if we hired her." Ellie said. "There has been some negativity around her." She says that the teacher has been around students when they were not behaving and allowing it. A teacher wanted to correct the students but didn't want to undermine the substitute's authority. "They were with an adult" is how Ellie phrased what the teacher had said. Ellie concluded saying, "There's a felt perception that she's not competent."

When asked to evaluate the candidate, the three questions that come up immediately and organically are: performance, trajectory, and philosophy. The candidate passed the performance test, as the school's director said "she did a nice job." She demonstrated the potential for growth and development by being reflexive about her own shortcomings as a teacher. However, her

reported disbelief in no-excuses schooling was a red flag. While she was committed to the school, she appeared to question its foundation.

Finally, other faculty questioned the teacher's ability to discipline students. In this case, the assistant was seen talking with students in the cafeteria when those students should have been returning to class. Her reputation in general was that she was "chummy" with students, a better friend than disciplinarian. Though, in my experience, she was as quick to fault students for being undisciplined in conversations with other teachers and wanted to ratchet up punishments for students at the end of the year when the school began to crack down on misbehavior.

She was ultimately not promoted.

It is possible that, had the teacher been a young white woman, her talking with other students may have been read by other teachers as warmth and breaking rules to have a conversation with students may have been interpreted as a legitimate use of discretion rather than as being chummy. I don't have the data to make that claim. What this does show is that empathy with students and warm relationships played no role in evaluating this teacher. Instead, she was held against the ideal type – the true believer capable of sweating the small stuff and developing into the kind of teacher the school wanted.

The school succeeded in diversifying its new hires. Of the five full-time teachers hired in 2015, none were young white women and only one was a graduate of NETT. They did not relax their standards but reinforced them. If this push for diversity had represented a more clearly articulated strategy, the values embodied by diversity would have been incorporated into the notion of the ideal teacher. Instead, they sought type-A teachers with different experience levels, genders, and races.

PROFESSIONAL COMMUNITY: A PRESSURE COOKER FOR HIGH ACHIEVERS

Teachers at charter schools are expected to work longer hours and longer school years than teachers at traditional public schools. Moreover, the work is taken to be harder because students from low-income families are more challenging and, having fewer students, the school has fewer resources. These working conditions by themselves do not predict all of the turnover at charter schools. What is missing from these accounts and what I find at Achievement the professional community ties three factors together to create a high-pressure environment among teachers: a high work load, a tightly coupled community, and high expectations for performance.

Not only are charter schools known for expecting a lot of their teachers, but they are also known for their focus on on-the-job training and increased duties for teachers (Furgeson et al. 2012). On the job training and extra duties come on top of the traditional expectations for teaching. To get it all done, teachers could shirk their responsibilities. However, this work is embedded in teachers' social relationships, whether meetings, observations, or collaborations, which hold them accountable. Embedding extensive responsibilities into teachers' social networks produces a tightly-coupled community. To manage the demands and oversight, staff could lower their quality of work – get it all done and turned in but at a lower quality. Contravening this however are the high expectations for teacher's work. Not only are teachers asked to do a lot and always observed, but the goal of observation is to ensure high performance.

In other words, the professional community forced them to be true-believing, type-A achievers.

Extensive Duties

Teachers at Achievement were asked to do a lot in their daily job. Each full-time teacher taught 4 hours per day – two, two-hour courses in middle school or four 50-minute courses in high school. They also were responsible for roughly an hour of duties per day, such as monitoring lunch, detention, or the reading lab. Each week, they were expected to attend a one to three hour faculty meeting, an hour long department meeting, hour long grade level meeting, hour long one-on-one meeting with their department chair, an hour-long meeting with all faculty in their school; a 45 minute community meeting, and a 30 minute advisory meeting. That comes to about 30 hours per week on average in which teachers are required to be somewhere doing something.

On top of that was added the daily, open-ended work of lesson planning and grading. Teachers were expected to give grades four times a week. Two assignments were to be graded for participation – did students do the work? – while the other two were supposed to be graded for accuracy. Teachers were expected to keep these grades in their gradebooks checked by their chair each week as well as an electronic grading system maintained by the school.

Lessons plans were expected to be scheduled out two weeks in advance with unit tests, midterms, and final exams written out at the beginning of the year. "Backwards planning" as it is called, helps teachers work backwards from what they want students to know on their tests to the daily lesson plans that will teach those material. However, this meant that when plans change, whether because students did not understand the material or an activity failed, everything forward had to be reorganized – ten lesson plans, the unit test, the midterm and final. Furthermore, teachers course assignments were regularly changed. During my field observation, no teachers had been teaching a class longer than four years and the mean tenure in a class was

just under two years even though the average teacher tenure was closer to four years. Despite the fact that many teachers kept copious documentation for every class less, every time a teacher teaches a new class, their prior year's lesson plans became useless. (New teachers reuse about 25% of the lesson plans from the previous teacher.)

Finally, there was a bevy of irregular work like calling parents with updates (a formal requirement of the job), entering information like attendance, discipline, and grades into the school's tracking system, meeting with parents and students, overseeing clubs, chaperoning field trips, or attending after-school events. Teachers are expected to give students their personal cell phone number for homework helping, ensuring teachers are never truly off the clock.

Tightly Coupled Community

Teachers are held accountable to these work requirements through regular oversight. Teachers are observed weekly by their chairs and other teachers, school administrators, and visitors. Teachers meet weekly with department chairs and occasionally the dean of instruction to review their teaching skill, lesson plans, and grades. The weekly or biweekly grade level meetings hold them to account for student behavior, duties, and participating in school events. The point of most of these observations meeting is as much about teacher support as they are about surveillance and monitoring.

Beyond the weekly accountability meetings, many meetings serve the purpose of development – helping teachers improve their craft. There is an all staff meeting every week and semi-regular high school and middle school meetings for professional development training, data review, policy discussions, and school announcements. There are annual periods of Instructional

Roundtables where teachers collaborate across departments for instructional development and to visit other schools to learn from them.

This intertwining of duties and social interdependency creates a tightly coupled organization when teachers and staff hold one another accountable to a heavier than average workload. This model of regular observation and collaboration contrasts sharply with traditional image (and reality) of public school teaching (Lortie 1975; Sarason 1971). In most schools, teachers work in isolation. They are observed as much as twice per year for their annual performance review and their classrooms are rarely visited by other teachers. When teachers want to grow as professionals, they must take time off and go elsewhere for formal trainings by themselves.

This collaborative approach to teaching has been championed by education scholars as a professional community. In their review of the concept of professional community, Bryk, Camburn, and Louis (1999, 753) characterize it as "schools in which interaction among teachers is frequent and teachers' actions are governed by shared norms focused on the practice and improvement of teaching and learning." Bryk et al. then list three essential features of a professional community: "(a) reflective dialogue among teachers about instructional practices and student learning; (b) a deprivatization of practice in which teachers observe each other's practices and joint problem solving is modal; and (c) peer collaboration in which teachers engage in actual shared work."

Here I find that there is a dark side to this professional community – it's driving up teachers' work at Achievement. Whereas hours in professional development used to be spent over the summer or during teacher training days, professional community is shifting that work to the school day itself. Teachers are still doing the formal, distant training over the summer, but at

Achievement College Prep, they are also spending up to five hours each week in development. Furthermore, by interweaving responsibilities across teams, everything seems important all the time. When teachers fell behind, department chairs cut teachers some slack on their grade books and lesson plans. But this didn't translate into forgiveness by grade level leaders when asking teachers to pitch in for school events, calling parents, or logging student misbehavior.

Addressing Overwork

It was apparent to everyone I asked that teachers were overworked. Each was their own case study. When I asked interviewees how much time they worked, they reported between 60 and 80 hours per week. When I asked if teachers worked too much in general, staff would point to the same kind of story of teachers coming in over the weekends or working at the school until 2 or 4 A.M.

In response to the problem of overwork, the school's director implemented several practice changes when she took over. She locked the school at night and on weekends as a symbolic act (since teachers have keys to the school) to encourage teachers to go home. She announced a policy in which no one was expected to answer emails on weekends. Finally, the school implemented a formal work-flexibility policy wherein teachers did not have to be in the school during school hours if they did not have to be there. Whereas all teachers had been required to be in the school building from 7:30 until 4:00, with the new policy, if teachers did not have a class, duty, or meeting until 9 A.M, they did not have to be in the school until 9 A.M.

These changes did not address the underlying work requirements, but rather focused on loosening oversight. Teachers still came in on nights and weekends. The email policy, announced at the end of the 2013-2014 school year, was largely forgotten by the end of the next

school year. While the work flexibility policy was considered a feather in their cap, particularly in recruiting new teachers, teachers worked the same hours from the comfort of home.

In interviews and observations, no one pointed the finger at the amount of work teachers were expected to do as the cause of overwork. And there was no discussion of changing the work expectations of teachers as a way to address overwork. Instead, the change the school sought during my year in the field was focused on providing more structure for teachers' development.

The Ideal Type in Action

The effect of the school's professional community was to maximize workers hours. This was not the administrations' intent however. Rather, they created a system that played to teachers' need for achievement and perfectionism. Teachers, like the school, were always expected to become better and more effective and to work relentlessly towards their own improvement. The system they were creating was built to ensure there was always a problem for teachers to fix or a weakness to bolster. There was always a way to be better.

During my focal year, I had the opportunity to witness this firsthand as the school's leaders redesigned their teacher evaluation rubric. When I arrived at the school, the goal for teachers' professional development was not entirely clear despite all the work put into professional development. The school's administrators defined good teaching as "masterful instruction," which meant teaching rigorous content and maintaining good student culture in class.

This model of masterful instruction was only loosely connected to teachers' development. For example, department chairs set developmental goals for each teacher. These goals would (ideally) be short, definable goals like having a teacher wait longer after asking a question to

give more students a chance to answer on their own (what is called "wait time"). These goals would last several months and, when a teacher improved, the chair would replace them with new goals. But these goals were only loosely connected to "rigorous content" and "good classroom culture" and were defined helter-skelter.

In an interview, one administrator described the disconnect this way:

> I think one of the challenges is assessing, "what does the goal mean in a larger framework of how a teacher is progressing in their career? Are they making progress early enough. Is this the right goal at the right time?" And while there was conversation about the goals by well-intentioned, smart people early on, without that clear road map of a rubric, I think it's hard for teachers and department chairs to get as much leverage out of the goal as we might want. Whereas a teacher can say, if I clear this goal, I'll be in the three column instead of the two column. It's more concrete that way. So, there's a little bit of floating out there, "this is my goal, I work on it" but what does that lead to?

During my focal year, the administration worked on such a framework to link everyday, achievable goals with a broad vision for masterful instruction.

The Academics Committee was charged with creating the framework. Made up of the department chairs and school's senior academic leadership, the committee had also overseen the original implementation of Masterful Instruction and guided the process of setting goals for teachers. Committee members began by using a teacher evaluation scheme developed by The New Teacher Project (TNTP) as a template. In its original version, the TNTP evaluation scheme gave teachers a numeric rating of one to five on up to five items in five topic areas: planning, course content, classroom culture, the environment, and assessing students' learning. The primary discussions among the Academics Committee were how many points teachers should be given, eventually settling on a scale of 1 to 5; which of the two-dozen measures in the five areas

should be scored; how to rate different kinds of in-class observations; and what and how to communicate observations and evaluations.

The resulting evaluation scale promised a comprehensive trajectory for all teachers. It measured all of the pieces viewed by the Academics Committee as constitutive of good teaching and could apply to all teachers from middle school math to high school art. The idea that teachers could achieve perfection by getting fives on all two-dozen items was not seriously considered. Instead, the assumption was that teachers' weaknesses would show up among the items and the evaluation scheme would map out the path forward. There would always be a problem to fix or a weakness to bolster.

As one teacher described it, the high expectations and tight oversight ensured that every weakness would be

> Interviewer: So the question is why do people work so long?
>
> Interviewee: I don't know a hundred percent. I think it has to do with wanting to be perfect.
>
> [Here it's] like, "oh, you're teaching that? You need to fix that. Let's stay after school, fix that. We'll meet tomorrow at 6:30[am]. We'll go over next week's lesson plans too to make sure you fix that element in all your lesson plans for the next week and then we'll follow up in a week." Which is great for a year or two. But people can't handle that.
>
> What happens is we hire super type-A people with super high standards for themselves and then we hold them to really high standards and identify every little mistake they make. And there's no acknowledgement that imperfection is part of life. If it's not perfect, you didn't try hard enough to make it perfect.
>
> I could show you the emails, [my chair] sent me an email at 12:22 am. I'm not immune to this. I'll do all this stuff. But she'll send me stuff…people send emails at 5:15 in the morning. So you get these things. And I'm the kind of person who will do that. But I'm also the kind of person where, if I'm not surrounded by it, can probably ease off. But when I'm surrounded by it, I will not ease off. I'll be a total nutso. I'll put my physical and mental health and my relationship with my partner in jeopardy.

RETAINING THE BEST AND BRIGHTEST

In addition to hiring under-prepared teachers susceptible to overwork and burnout and building a pressure cooker community, teachers at Achievement also work with job insecurity, for less pay, and no advancement. By job insecurity, I mean that teachers stood to lose their jobs on a regular basis. This was the case because, being on annual contracts, the contract renewal processes routinized the question of whether or not a teacher should be fired. For compensation, teachers at Achievement started with relatively competitive salaries, but only received cost of living raises over time. They quickly fell behind their traditional public-school peers who received pay raises each year for their work experience. Finally, there was no predictable path for advancement at Achievement. Many teachers interpreted this as meaning they had no future at the school. Each of these was justified by the same beliefs about teachers that led to an underprepared and overworked staff.

Teacher Evaluation and Contract Renewal

Teachers at Achievement were hired on annual contracts. Each spring every teacher had to pass their evaluation to be offered a new contract for the following year. Despite the frequency and ubiquity of this routine, few teachers were "not asked back." In my focal year, only one full time teacher was not offered a new contract. The other six full time teachers who did not return did so at their own discretion. The effect of the contract renewal routine then was not to fire under-performing teachers, but to provide stakes for the schools' expectations – teachers who do not do the work and who do not get better every year would be let go.

The most commonly discussed reason for a teacher not being asked back was for a teacher not improving enough during the year. As mentioned in the previous section, department

chairs set improvement goals for faculty. Chairs then evaluate teachers on these goals in their weekly observations and provide feedback to teachers on how to continue improving. Chairs track these goals across the year. When a teacher was deemed to have accomplished their goals or needed a new goal, the chair came up with the new goal. This system ensured that, no matter the seniority or experience of the teacher, there was always something for a teacher to work on and even senior teachers could be fired for inadequate development.

This growth-based evaluation was not designed by the school but evolved as a byproduct of shifting responsibilities among managers. The annual evaluation used to be more like that found in traditional schools. Principals would perform a few observations and then make a hiring and firing decision through what was described to me as a "good cop, bad cop" scenario. Principals and the director would evaluate teachers while department chairs acted as advocates for a teacher. However, as one administrator put it in an interview, "It set up a really funny dynamic where the person evaluating you is the person who's in your room least often. And it also limited the impact that department chairs were having." The problem was that contract renewal decisions were made with less information and the chair, who was tasked with enforcing academic policy and running academic programs, had little real power over their teachers.

The logic of growth emerged in solving these problems.

> Years ago, [teacher evaluation] was more traditionally based. You got 2 observations a year. They were formal, you knew they were coming, and occasionally there'd be some walkthroughs. Like 4 or 5 years ago, it shifted to where department chairs were coming in every week.
>
> I think maybe three years ago is when we shifted to having department chairs as a part of the evaluative process. Which is a change and it's interesting because you're a coach and you're evaluating. But with that, the focus has been (and hopefully it's transparent to people) that when we step in the classroom, we're not evaluating you. What we're trying to do is coach and develop and what we're evaluating is how readily available

> you are to the feedback. How open are you to receiving these suggestions? Where are you in incorporating these into your practice? Not, is it good when I walk in?

The evaluation system focusing on growth emerged over a series of changes. First, teacher oversight sifted from biannual evaluations by a senior administrator to weekly observations by the chair in addition to the biannual evaluations. The information asymmetry this created between chairs and administrators was resolved by bringing chairs into the evaluation process. In doing so, the standards for evaluating teachers shifted from the quality of classroom teaching as observed twice a year, to the teachers' ability to grow on a weekly basis.

Teachers internalized this grow-or-go attitude. As one teacher said, "if you're struggling, you might get fired because you're not doing your job. The fact that we have a contract where you might be offered a place back, maybe not is kind of like, if you improve this much you get a contract, if not, goodbye." When I asked Ellie, the chair of the teacher who was not asked back, how the decision had been made, she was open to talk but careful in her words. She pointed to the teacher's "trajectory." The teacher was not as expert as their years of experience would imply and wasn't exhibiting the growth over the year they expected of someone with their experience.

Contract renewal was an annual ritual central to teachers' lives. Yet, the school worked to minimize its significance. For nearly all teachers, contract renewal was portrayed as a formality, "you have nothing to worry about," or as an affirming experience "we think you're a good teacher, we want you here." Teachers on the cusp of not being asked back already know ahead of time where they stand because of the remediation steps the school takes well before the annual review. For the rare teacher who isn't asked back, the message is they have not met the expectations for someone at their experience. The effect of the annual review and contract offer is to reinforce the grow-or-go model.

Compensation

Despite the fact that teachers who were asked back were supposedly getting better every year, their compensation remained flat – all full-time teachers at Achievement were paid the same. Compared to district schools, teachers were hired at roughly the same rate. However, wages quickly diverged as the salaries for teachers in traditional public schools grew annually. But for teachers at Achievement, the only pay increase was a cost of living increase of 2-4%. By the time a teacher was in their sixth year, they made $20,000 less than a teacher with equivalent experience in the local district.

The flat pay structure was based on a contract logic – teachers were paid for performing a predefined set of services; four classes, two duties, etc. If they performed those services effectively, they were hired again. If they wanted more money or more time off, they had to move into different roles. For example, assistant, substitute, and part-time teachers were paid less and had fewer responsibilities. On the other hand, department chairs, grade level leaders, and administrators are paid more but substituted managerial responsibilities for teaching.

When the new director was hired, the board had identified teacher retention as one of her priorities. They were losing teachers in their third, fourth, and fifth years. The year before I began my observations, the school implemented a signing bonus for teachers who signed new contracts in their third, fourth, and fifth year. These bonuses went away starting in the sixth year. In an interview, a senior administrator explained the rationale saying, “We’re trying to find little ways to get people to come back. And then, the hypothesis, is that once they’re in their sixth year, they’re staying in the organization for different reasons [other than money] … they’re not potentially as mobile.” Another senior administrator clarified the “different reasons”;

> I think that [teachers] know they're doing the work and getting paid a little less than teachers in other schools, but knowing that at the end of the day 100% of our kids are matriculating to college. I think there's the idea that... the culture is that we're scrappy and happy...would be maybe a way to characterize it. And when you're scrappy, things can get tense because you're working on a razor's edge of time or a budget or exhaustion or personal sacrifices. It's not always going to be crazy joyful. But at the end of the day, and you could ask any staff member, they would say they're mission driven and if you asked any kid why they're here, they would also point to the mission. Like, being 100% bought into why they're here and why we're here -- that everyone is charging toward the mission together.

While the original system of compensation reflected a logic of contract work, the new system reflected the logic of the calling – that one's true rewards were transcendental and immaterial rewards of being part of something meaningful. In all of this, the logic of paying the best and brightest teachers a premium for being more productive than less effective teachers, more valuable to the organization, or more mobile were never mentioned. The logic of market competition driving good teachers into charter schools and bad teachers out was never brought up regarding teacher compensation.

Teacher compensation was not set as part of a comprehensive teacher hiring and retention strategy. Instead, it evolved organically in response to problems. When teacher retention was put on the table as a problem to resolve, the ideal type reared its head. Instead of providing competitive compensation, the school decided to pay just enough until they could leverage teachers' willingness to work for meaningfulness rather than money.

Advancement

The belief that teachers work for meaning also propped up the school's approach to career advancement. Despite the fact that career ladders are important for workers (D. C. Feldman and Ng 2007), there has been little recent work in education on how policies regarding

the structure of teachers' careers fits into teacher retention or performance (see Dee and Keys (2004) for an exception). At Achievement, there was no career ladder for teachers. As stated by the senior administrator earlier, the assumption was that teachers would want to continue teaching their four classes a year until retirement because they want to be part of the school's mission. This assumption began to be questioned at the end of my focal year when the evaluation rubric was being discussed.

During a faculty meeting near the end of my focal year, the school's director gave staff an update on the teacher evaluation rubric and, for the first time, mentioned teacher advancement. She tied the evaluation rubric to the new signing bonuses as something teachers had been asking for, "People have said, we'd really like to have a career trajectory, something that ties what we're doing in to compensation. We recognized that we don't have a career trajectory…" In attempting to solve the problem of "what is good teaching," they began to see teachers' demands for a career trajectory and pay scale in a new light.

Teachers had long wanted a career ladder, though they never articulated it to me in those terms. One teacher who was leaving after their second year, told me that they couldn't see the benefit of staying at the school for another two years, "If I would have stayed here for four years, I would have four years of teaching under my belt, but I don't know what else would have come with it…besides teaching here for four years." This teacher was committed to the mission of the school. They were leaving to pursue education in an adjacent field. However, they did not see anything to show for that commitment. Other teachers wanted to stay at the school and continue teaching, but wanted to do more than just teach. For them, there simply weren't any options besides becoming a grade-level leader or department chair.

The intersection of the school's leadership taking on the evaluation rubric and teachers

ongoing demand for advancement in pay and responsibilities sparked a re-assessment of the school's career ladder. The academic committee did not come up with a solution by the end of the school year, but they had defined the characteristics of a solution.

In the same presentation,

> The director projects an image of the career trajectory for cooks as an example of a career pathway. It's byzantine with many intermediate paths from line cook to head chef. "We've started thinking about, what is the analogous pathway for teaching and when do we jump on and off the pathway. And, one goal was to make the pathway diverse. If you want to jump off the pathway into another function or to leadership you could or if you want to stay on the teacher pathway for your entire career you could."

The first characteristic for a successful ladder was diversity – giving staff options to switch between teacher, leadership, or functional advancement. It is important to note that, while the primary audience for the policy change was teachers, the ambition was for the evaluation and ladders to extend to all staff. Comprehensiveness was thus a second characteristic.

The third was growth-orientation:

> "We've been talking about, in drafting our model, there would be different levels. One thing that we have a difficult time with is that, if you're new to ACP but not new to teaching, we still treat you the same. We want to build in different levels where, you could enter [the school] at different levels. And, you might be able to move through these levels at different rates. The idea of a trajectory that acknowledges different ways of moving through the trajectory also acknowledges that you can move at different speeds.
> "We're also thinking about rewards like changes in compensation, other lines of work and opportunities.

At the heart of the new model was the belief that advancement levels would correspond to expertise and skill (not tenure). Rather than rewarding staff for doing a job, they would reward staff for being better or worse at doing that job. That is, the structure was built to encourage and reward personal improvement.

The ladder the school began to envision was one that would enable career and compensation changes to map onto and enable personal growth, whether that was learning something new or getting better at one's current role. The demands of teachers for more diverse careers tied to increased compensation, when combined with the problem-solving process of the new teacher evaluation rubric, led to the creation a structure that was meant to enshrine the growth ideal. Their commitment to the ideal type led them to create structures that would encourage teachers to fit the type.

Summary

The incentive system to stay at Achievement College Prep year in and year out is based on rewarding teachers with meaningfulness and never-ending growth. These systems were not the result of systematic planning but emerged from solving disparate problems around teacher retention. First, every year, teachers were fired or retained based on whether they demonstrated effective teaching, but it was unclear what effective teaching was and who should be making that decision. In restructuring the teacher evaluation scale and process, the school designed an evaluation system around the idea that all teachers whether novice or expert would grow-or-go. Teachers were paid less and worked longer hours. In response, they implemented signing bonuses for teachers most likely to leave to keep them around until teachers were more invested in the mission of the school itself. Teachers left because they saw no future for themselves at the school. To solve this problem, they began creating a career ladder rewarding teachers' improvement in skills and providing opportunities for growing in other ways beyond just teaching. The result was a structure that not only reflected the ideal type of teacher, but was meant to mold teachers to it.

WHAT'S GOING ON?

The churn and burn system at Achievement is a case in which notions of the ideal employee as one who is already emotionally and intellectually committed to the organization allowed them to build an employment system optimized to burn teachers out. By hiring, professionalizing, and rewarding teachers committed to the organization, Achievement created otherwise unsustainable work requirements, wages, and expectations.

This wasn't consciously planned. The lack of awareness about burn and churn is apparent in the failed attempts to mitigate it. The changes they attempted – closing the school overnight, not requiring teachers to be on campus, and not expecting teachers to answer emails over the weekend – address symptoms of the problem rather than its cause. Teachers weren't working late because the lights were on. They worked late because they felt guilty for not pulling their weight. They worked weekends because they couldn't create effective lesson plans during the week.

Despite the substantial amount of research on the working conditions in charter schools, there has been little investigation into why these conditions have prevailed. Only in commentaries can you occasionally find any attempt to attribute a cause to the churn and burn system or teachers' working conditions in charter schools in general.

Researchers in organizations and education who study organizational strategy view it as the result of a purposeful decision-making process whereby organizational actors define the organization's goals and select the means by which they can be best achieved (Cyert and March 1992). The churn and burn employment system among charter schools reflects neither the self-awareness of strategic planning nor the utility-maximizing assumption of rational decision-making. No one sat down and said, "schools work best when they burn teachers out, so what's the best way to burn teachers out?" Instead, churn and burn represents an odd category of

organizational behavior that's neither rational nor utilitarian. It's an emergent behavior that undermines the organization's goals.

The Logic of Cultural Fit

Accounts of organizational systems tend to portray them as the result of deliberative processes, extensive research, and systematic testing. Neo-Institutionalists view them differently, arguing that many systems come from taken-for-granted cognitive schema imported from the organizations, regulators, and commentators around them. Organizations import strategies because they confer reputational benefits from the environment rather than because they meet some technical requirement or actually improve organizational performance (DiMaggio and Powell 1983; Meyer and Rowan 2006).

One such imported system is an institutional logics. Institutional logics are ways of seeing the world that frame the way organizations strategize (Thornton, Ocasio, and Lounsbury 2012). They define the terms of the strategy-making process – should we reduce costs and maximize output or should we put the right policies in place. They also enable sets of practices and beliefs to "make sense" together. Organizations that see themselves as market actors maximizing profit, focus on cutting costs, maximizing productivity, and rewarding investors.

In the 1970s and 80s, the rise of service work led to a sea change in the way employees were understood. Service work required employees who could perform new kinds of social and emotional work – making customers feel comfortable, important, or welcome and managing one's own feelings about the work. We wanted employees who could not only meet the technical requirements of the job, but its emotional requirements as well (Hochschild 1983; Wharton 2009). Achievement College Prep needs teachers who will be *disappointed* when students make

mistakes, *angry* when other teachers slack off, and *earnest* when articulating the vision of the school. One can either train teachers to feel these things in the appropriate situations or, better yet, hire employees already capable of them, who already "fit" the culture of the organization.

The logic of cultural fit "makes sense," but has been criticized recently. When hiring new employees, "cultural fit" has been second only to technical skill for tech company hiring managers. Companies not only wanted to hire the best coders, but those who would get along in the company's culture. While the idea sounds right, in practice, fit meant hiring people who look like those who are already at the company – almost always young, white, men. Women, people of color, and candidates with other minority identities were seen as unlikely to mesh with the company culture. This filtered out talented minorities and led to ethnically homogenous companies and teams which have long been known to be less productive, creative, and successful.

The logic of fit worked differently at Achievement. The "fit" to be made was not among teachers, but between the teacher and the mission of the school. The ideal teacher was driven to close the achievement gap by tolerating no excuses. This meant three things. Teachers would be intrinsically motivated. Their reward is student success, not awards, wage increases, or time off. Second, teachers would attribute outcomes to their own actions. If the teacher does their job, students will succeed. This made teachers perfectionists. Finally, they were to be true believers. They believe in the no-excuses charter school's ability to end educational inequality in the United States and, by their own effort, can bring about social equality.

In the world of service work, it "makes sense" that employers want to hire people who want to do their best to achieve the organizations' goals because they believe in the mission. At the same time, one can see how this worldview might lead to overwork, under-payment, and high

turnover. But the logic alone is not enough. Churn and burn wasn't destined to happen. And it happened despite many attempts by the administration to stop it. How could such a problematic employment system arise without the express direction of school leadership?

Fake it until you Make it: Organized Anarchies and Emergent Systems

Scholars have investigated a variety of cases in which organizational systems formed outside of deliberate decision-making (Mintzberg and Waters 1985; Vaughan 1999). This either happens in uncontrolled; unreflective; or emergent systems. A system can become uncontrolled through accidents (Perrow 2011; Turner 1976; Vaughan 1997), errors (Ramanujam and Goodman 2003), or exogenous forces (Haveman 1992; Tushman and Anderson 1986; Zucker 1987). They can be created without reflection from leaders (Haveman 1993), local models (Meseguer 2004), and experts (Kelly and Dobbin 1998).

Churn and burn emerged across many disparate rational and reflective decisions (Andersen and Nielsen 2009). Emergent phenomena can be distinguished by five features: radical novelty, coherence, dynamic, recognizable and exist at a macro level (Goldstein 1999). We can see churn and burn reflect each. It is novel within education and recognizable and coherent enough to be given a name. As I found at Achievement, churn and burn developed over time in different parts of the school, making it dynamical, and these changes were interlocked, making sense only when taken together (i.e. they represent a macrosocial phenomenon).

Systems emerge when control over planning and execution are devolved to separate decision-making processes, none of which can wholly determine the outcome (Mintzberg and Waters 1985). It can happen when subunits influence collective strategy (Burgelman 1983; Noda and Bower 1996) or when leaders allow subordinates to formulate their own strategies (Bigley

and Roberts 2001; Child 1997). In this case, the system emerged as the logic of fit was used to solve a set of related problems – hiring, developing, and retaining teachers. They hired teachers who would fit, developed to teachers to be hard-charging true believers, created a system where they were always charging to the mission, and traded intrinsic rewards for wages.

In addition, the circuit breakers had been cut off. Based on what we know about burnout, value congruence short-circuits the connection between working conditions and employee engagement (Leiter and Maslach 2003). Churn and burn should collapse on itself. But people are willing to put up with a lot of crap to accomplish what they value. Because the school was successful in hiring and creating teachers who fit, it allowed the school to put more on teachers than it should have.

SOLUTIONS

It's indicative that the rise of emotional labor in service work coincided with the growing recognition of burnout as a problem in human service occupations like teaching (Cherniss 1980; Freudenberger 1975). What made burnout unique at the time was that it was characterized by an emotional and cognitive exhaustion. Rather than being physically exhausted or injured, employees experienced emotional exhaustion, cynicism and empathic disconnection from clients and coworkers, and the loss of a sense of efficacy (Maslach, Schaufeli, and Leiter 2001). The entwining of emotions, meaning, and work made our humanity subject to consumption.

This doesn't mean we should stop telling teachers to do what you love or hiring teachers who believe in the work they do. The problem is not our passion. The problem is that passion gives leaders and decision-makers a false sense of organizational capacity – that people or

departments can do more than they really can. In an organization where people are true believers, extra steps must to be taken to ascertain an organization's true capacity.

Further, had a different definition of the ideal teacher prevailed at Achievement, a different system would have emerged. If the ideal teacher was seen as a person with healthy boundaries, it is doubtful they would have favored younger teachers. If the ideal teacher was seen as a person who takes ownership over their own development, the professional community would not have been so tightly connected. There are other ways to change the individual components of the churn and burn system, but a change in the logic of the ideal teacher would change them all.

Finally, what gives emergent systems their shape is the repeated but disconnected solving of problems. Cohen, March, and Olsen called this form of decision-making the "garbage can" (Cohen, March, and Olsen 1972). In the garbage can, problems, people, and solutions get tossed around the organization. What problems get solved by whom and how is a matter of circumstance. Systems emerge when people, problems, and solutions start clumping together. In this view, churn and burn would be different if the school had attempted to solve all of these problems at once. If the school's leaders had sat down and come up with a system linking hiring to training, and promotion; the problems with the system would have become apparent.

A solution that is unlikely to work is importing an employment system from outside the school. For example, no-excuses charter schools could have their teachers join the local teachers' union. That would likely end or substantially reduce many aspects of churn and burn. This is infeasible because it doesn't fit with the school's vision of the ideal teacher. Even though I observed no anti-unionist beliefs unionized teachers are generally not viewed as all-in on student

performance by charter school advocates. For an existing employment system to be brought in from the outside, it must be consistent with the logic of the ideal teacher.

CONCLUSION

Emergent systems are the unknown unknowns of the organizational world. They appear out of nowhere and without explanation. They are "just the way things are." As policy makers, how do we build organizations and schools that are resilient against developing cancerous systems, where what emerges is positive? As researchers, how do we identify the levers to control processes whose shape is only apparent after the fact?

First, we need to understand what interventions affect emergent processes. Complexity theory approaches interventions indirectly. With emergent processes, the challenge is that the mechanisms and systems leading to the outcome are too varied to understand, model, and therefore control deterministically (Anderson 1999). Whether there is such a thing as an irreducibly complex system in a metaphysical sense is irrelevant. We simply do not have the resources or cannot devote the resources to know the system and thus cannot specify the steps needed to fix it. Thus, emergence requires a non-mechanistic mode of causal explanation.

Rather than focusing on if/then models of deterministic control, emergent processes can only be enabled or shaped by providing systems with the resources, rules, and contexts they need to achieve some desired end state (Anderson 1999). This could mean changing the demography of the organization (Lawrence 1997), its incentives (Lewin, Long, and Carroll 1999), or its structure (Meyer, Frost, and Weick 1998). When looking to build strategy around processes that are unclear or uncontrolled, we should look to the "ingredients" for success.

In organizational theory, there are two ingredients-based approaches. One is the resource-based view of the organization which posits that organizations succeed by possessing the most strategically valuable resources (Kraaijenbrink, Spender, and Groen 2010). While this line of work has been focused on the question of organizational survival in competitive landscapes, it is possible to adapt it to the more general problem of organizing. Specifically, this theory views organizations as driven by technical requirements, whether profit or student test scores. Organizations succeed based on their ability to acquire and use the most valuable resources to meet those requirements. The question for researchers and policymakers is what resources are needed to support the emergence of effective systems? Would more funding, better trained teachers, or access to support services like human resources or operations lead to better run schools?

Growing out of the resource-based view, dynamic capabilities theory posits a similar view of organizations as driven by technical requirements. But rather than competing on a static body of resources, dynamic capabilities sees organizations as succeeding through dynamism (Wang and Ahmed 2007). Organizations are always evolving and adapting, and the degree, direction, and success of that adaptation is determined by the organizations' people, processes, and routines accomplishing the adaptation. Rather than pouring resources into a problem, the organization with dynamic capabilities has a team or unit that can effectively decide what the right resources are or if more resources are even the right solution. For policy makers and school leaders, the question is what kinds of teams, practices, or routines lead to better run schools?

In education, there is one line of work that takes this enablement approach to school effectiveness. Sebring et al (2006) argue five characteristics lead to school improvement over time – leadership, parent and community ties, student-centered learning, ambitious instruction,

and professional capacity. Schools that are strong on these five characteristics improved their test scores much more than those that were weak. While this is far from the only study arguing for direct relationships between organizational factors and student outcomes, the model is not mechanistic. As they say "our framework suggests a dynamic model of improvement, not a specific lockstep plan. Progress can advance along numerous paths, and no one course is obviously best for all schools … sustained work must eventually emerge on each of the essential supports" (Sebring et al. 2006, 16).

The problem with these theories is that they are focused on organizations' core goals and their core competencies. Sebring et al.'s research is grounded in student achievement. It provides no evidence for whether the five characteristics will help in areas outside of test scores. That is, we don't know if it would useful in addressing teacher pay, diversity, or civic engagement – the kinds of problems schools can face but where they often have less clarity and control. Similarly, the kind of adaptability championed by the dynamic capabilities perspective may lead to continual change and burnout.

The theories are not weak. They are untested in a large class of important cases. In general, education and organizational scholars do not study these secondary or supplemental issues. We investigate them piecemeal with a study on burnout, diversity, or social responsibility. In order to develop policies around enabling emergence, we need to examine behavior away from organization's core competencies as a *category* of phenomena. Organizations are not simply production machines, but micro-societies where people find meaning, receive benefits, and socialize. These peripheral goals of an organization must be studied, both for their own sake, but also because this is the domain of the unknown unknowns. We need to study the well-rounded organization.

CHAPTER 6

WHY CAN'T WE CREATE EFFECTIVE SCHOOLS?

THE AMBIVALENCE OF THE NO-EXCUSES REALITY

At the outset of this study, I sought to understand why we had not created effective schools in the past forty years by examining whether no-excuses charter schools might be the exception. These are big questions which no single study can answer in total. I sought to answer a very specific version of these questions – whether it is plausible that no-excuses charter schools might be more effective. While causal studies indicate these schools are more effective at teaching students more than their traditional school counterparts, there is a debate over whether it is because these schools keep students in class longer or whether they actually do anything different. Research suggests that these schools are more effective because they have strong school cultures, a focus on data and testing, and strong teacher development programs. The goal of this study was to look in detail at one high-performing charter school, to see if these three factors could explain their success.

The answer is that the impact of these factors is ambivalent. Achievement certainly had a strong culture, an intense focus on testing, and rigorous teacher development program. As no-excuses charter schools go, Achievement is an excellent case. Yet, during my focal year, each initiative faced severe problems that potentially undermine the benefits each one offers. It is beyond the scope of this study to identify whether the good outweighed the bad. That would require studying comparable schools and comparable years to see if these problems are endemic to the no-excuses model and if those who have solved these problems or experience worse versions of them perform any better or worse. Instead, this study lends plausibility to the

prevailing conclusion that these effects while likely positive in aggregate are highly variable and hard to pin down. In other words, the causal studies are not clear because the reality itself is unclear.

This study and discussion set aside questions about the other potential benefits and drawbacks of no-excuses charter schools besides learning. For example, Golann (2015) argues the stringent disciplinary system encourages rote compliance among students and stifles creativity and free thought, the skills students need in college and life. Students may learn more but graduate without the soft skills needed to navigate higher education and civic life. I rely on learning outcomes as the measure of school success because that's what "effective" has come to mean for better or for worse. Doubling the number of students sent out of class is a bad outcome because for the school because it is time students are not in class to learn.

If Achievement College Prep suggests that no-excuses schools innovations have ambivalent affects, what's the path forward for education? What have we learned about the past forty years of education reform that can help us make real improvements over the next forty years? I explore these questions here.

DREAMS OF REFORM

The predicate of school reforms launched after the publication of *A Nation at Risk* was that we could make schools better through science-based policy and practices. For no-excuses schools, research had shown it was a lack of good management and culture that left disadvantaged students behind. For the state, it was about tying accountability to research-based curricular standards. For school administrators, building the recommended professional

community was an integral part of the strategy for making more teachers effective. In this study, these initiatives failed because they were, in one way or another, impractical.

None of these models of change are as easy as they're made out to be and science is not set up to provide solutions. The problem is not complexity per se, but the practical realities of organizing are not taken into account when designing interventions. Good management is a great idea, but creating a culture is no easy task even for the well-trained in a ready and willing organization.

Reform is necessarily done at geographic, social, and idiographic distances which require simplification. From a stage in Washington, DC, Albert Shanker (1988) articulated the vision for charter schools that would find its way into almost every state in the country. A platoon of researchers examining network effects and in-school development programs distilled down their findings into the idea of a "professional community." Law makers require standards and that schools shall meet them as demonstrated by students scores on a standardized test. We cannot abandon these distances and simplifications without abandoning reform. Instead, we need to make reform more reflexive of the design and process challenges schools face.

As I mentioned at the outset, one of the problems with social science is that we make the leap from explanation to control too quickly. When we find an association between some factor like class size or culture has an effect on learning, we recommend schools reduce their class sizes and strengthen their culture without empirically supported ways of going about it. When we identify a problem, say low expectations or a lack of competition, we tell schools and districts to raise expectations and create a marketplace. If the empirical chapters in this dissertation have demonstrated anything, it is how these well-intended prescriptions inevitably run up against organizational processes that make them infeasible, ambivalent, or counter-productive. To build

effective schools we need the science behind reform to be tightly linked to the design of practicable interventions.

BUILDING EFFECTIVE CULTURES

Organizational culture is a quintessential best practice. Well-supported by empirical research, easy to simplify and distance, and evidently feasible; culture should be an area in which schools are able to improve in a predictable, controllable way. Yet, at Achievement, the one year I observed demonstrated the opposite. The culture, well-established and clearly strategized, still failed to deliver results.

Certain things did work well. By the time I arrived, there was already a well-entrenched foundational culture in the school on top of which warm-strict and the culture of consistency initiative could work. Teachers and staff were all attuned to the culture of no excuses and the norms, beliefs, and values that entailed (even if they didn't always live up to those expectations or forgot them at times). They talked about it in meetings, used it in their work, and interpreted one another's behavior according to the no-excuses culture. Even though the only data I have, the culture in 2014-2015, suggests the culture is a failure, the literature suggests the foundational culture should prove more effective than not having one.

However, culture is only as good as the one you achieve. That the particular iteration of the culture in the 2014-2015 school year was counter-performative shows that simply having a strong culture is not good enough. The correlation between talking about consistency and the number of sendouts suggests that the culture of consistency initiative itself drove the breakdown. And, while research suggests culture is important for performance, it has nothing to say about the effect of warm-strict culture or the culture of consistency initiative in particular. As reformers,

we may be able to say, culture is important and strong culture is positively related to learning. But we cannot say what kind of culture is positively related to learning. Yet, this level of operational specificity is exactly what is needed for this type of change.

Lesson 1: We must turn research into "pills" for schools

Culture stands as an exemplar of a particular type of strategic change called performative change: using scientific research and theory as the basis for creating the thing theorized. No-excuses charters as a whole are in some ways an experiment in the impact of culture on schools. In my focal year however, the faculty and staff at achievement were swamped by uncertainty. Performative change at Achievement came with a clear strategy and feasible system. But things did not go off without a hitch from the get-go. And, every error in the strategy served to make it more and more difficult to disentangle what was implemented well and failed because it was the wrong thing to do versus what failed because it wasn't implemented well. Whether the strategy was wrong or the execution poor is unknowable from the perspective of the school. My data sleuthing leads me to believe the culture counter-performed, but I cannot say whether it was because of the contradictions at the heart of the culture of consistency initiative or because teachers failed to apply it appropriately.

Performative change is the equivalent of a prescription for effective schools – take this pill twice a day and you'll cut send-outs in half. The problem with this type of change is that the path to success is infinitesimally narrow. Everything from the diagnosis of the problem to the analysis of the results must be correct for the intervention to work and for us to say we *know* that it worked (or didn't). At present, social scientists can say with great confidence that good school culture improves student behavior. But we have no research suggesting that implementing warm-

strict culture through teacher training at the beginning of the year cuts student disruption in half in eighty percent of schools. And yet, this level of specificity and predictability is exactly what schools need from us to know whether what they're doing is effective.

Like most performative changes, the problem with recommending good culture is that we're skipping a lot of steps when we try to turn a descriptive model, even a causal one, into a program. In the absence of vetted programs, organizations are forced to invent them as they go along. Not only does this bootstrapping lead to variation in programs and success but it makes it impossible for the organization to distinguish success from failure. If we want research to drive interventions, we need to do research on the interventions themselves. Practical recommendations are not enough. We also cannot forget to include outcome measurement, reporting, and review mechanisms in our design. For example, if you're taking a pill, there's a set dose, you visit the doctor regularly for checkups, they ask a standard battery of wellness questions, and if things are better or worse and depending on side effects and life style changes, you take more or less pills. There is an entire *treatment plan*.

BUILDING EFFECTIVE ACCOUNTABILITY

Achievement was the ideal school for education policymakers. They worked diligently to implement Common Core standards in their classes, understand the underlying logic of the standards, and prepare their students for the new testing format. They brought in coaches and consultants, sent teachers and staff to state trainings, and worked with a testing company to perform their own Common Core-aligned diagnostic tests. Yet, students' scores tanked on the test, though not the state's other standardized tests.

Accountability has been the government's primary lever for creating effective schools and, in the focal year, we can see one reason why it has largely failed to do so. The shift was well-intended. Common Core standards reflect research-based best practices – deep knowledge and skill development rather than broad knowledge and rote memorization. But it was unclear how teachers should actually go about developing skills and deep content knowledge. Using the metaphor of the marshmallow challenge, I argued that state governments set learning goals for schools without also providing ways schools can meet that goal. If we applied a similar accountability model to hospitals, it would be like asking hospitals to increase patient's cancer survival rates without regard for whether the underlying medicine has improved. There are two lessons to be learned.

Lesson 2: Tie accountability to impact-based practices

Regulation takes two forms – regulation of inputs and regulation of outputs. Requiring seat belts and airbags is an example of an input-based regulation. In schools, input-based regulations set the rules for the number of students served by a school with what resources. The crux of the U.S. Supreme Court's *Brown v. Board* ruling requiring school desegregation was based on unequal educational inputs. Due to compulsory schooling laws and budgetary autonomy, schools and districts have control over the number of students they serve and the resources devoted to each student. It is thus fair and reasonable to hold schools accountable if they do not serve enough students or provide enough resources per student.

Fuel efficiency standards for cars and test scores for children are both examples of outcome-based regulations. These regulations allow the regulated to do what they want so long as what they produce complies with some rules. For example, the state leaves open how teachers

should teach. They can lecture, flip the classroom, do workshops, or go to the zoo every day; so long as students demonstrate mastery of standards on the standardized test.

My observations reveal how unjust this model is. Regulation must be constrained by the means available to comply with the regulation. We can't ask automakers to increase automobiles' fuel efficiency to 100 miles per gallon of gasoline next year because the technology for electrifying vehicles is not yet available for most models of automobile. However, we *could* increase fleet fuel efficiency to 100 miles per gallon in 5-10 years because the technology for doing so and trade-offs involved are well-defined. This is not the case in education. Although the school was fully committed to meeting the regulation, they failed to do so. And it is impossible to know whether they could meet it in the next 5-10 years, especially now that the state has abandoned the PARCC test. The outcome-based regime in education ties accountability to something schools and teachers cannot effectively control. We do not know what technologies (whether actual electronic and computer technologies or innovative teaching or schooling practices) could lead to great gains in learning. It is unjust (and ineffective) to reward or punish teachers and schools for outcomes they cannot control.

Output-based accountability cannot work in a field like education where the tools of success are unreliable or unknown. Instead, we need impact-based accountability. If states want to create effective schools, we must first devote our efforts to identifying those most impactful inputs and rewarding schools that use them. Standardized testing should only be a diagnostic tool. But there should be much more of it.

Lesson 3: There is no education without testing

Regulation is a dynamic and reactive system. States are constantly changing their tests, standards, scoring methods, and accountability formula as research, priorities, and politics wax and wane. And, because these tests are used to allocate resources, schools, teachers, test-prep companies, and regulators are constantly adjusting their strategies to meet the evolving tests. This dance makes test scores a poor measure of education because what students are learning and the measures of what students are learning are always changing.

Education researchers and policymakers focus on NAEP and PISA scores because they have resisted the dynamism and reactivity plaguing state testing. They have changed relatively little over time and teachers and schools do not design their curricula to prepare students to take these tests because they are not held accountable for students' scores. They thus measure a (relatively) constant form of learning independent of attempts to teach to the test. (Though, there are ways of gaming the scores by limiting the kinds of students who participate in PISA or NAEP.)

There are no such things as tests of what students have learned because there is no such thing as a perfect test or true learning to be measured. Students learn all kinds of things throughout the year and not just facts, new perspectives, or how to read and interpret a work of poetry. They learn how to get along with others, how to debate an idea, and when to challenge an idea. They learn things about sports, music, and their city – worthwhile things that aren't assessed on tests. There is an adage that a perfect map of the United States would be the size of the country itself. Similarly, testing, and any assessment of learning, is necessarily a partial test of specific things learned. It could be that American children are learning more now than they

ever have before. Schools may be teaching more than ever. We just are not testing what students are actually learning.

This argument is facetious but demonstrates the point that what we consider education is not learning, but the specific knowledge and skills we assess and the way in which we assess them. A corollary of this is that there are no effective schools apart from how learning is measured. This is true in two senses. As the saying goes, "if you're not counted, you don't count." If we don't test something, say civics or arts education, we can't say we're effective at it. Individual teachers and schools assess learning in these courses. But in lieu of standardized aggregate data, researchers and policymakers cannot say what students are learning or whether they are learning more of it now than before.

Second, because tests and teaching change, when measures of education change, the schools we deem effective can also change. The studies of charter school effectiveness have depended on state standardized tests. While this is a technical necessity, the results here suggest these gains may not generalize to other tests like the SAT, AP, or PARCC. The fact that there is no perfect test of true learning does not mean we should abandon testing. It means we should be much clearer on which tests matter and how those tests should be used for students, teachers, schools, policymakers, and researchers.

Some may argue that, while scores change depending on the test, education is not arbitrarily dependent on testing. Some students and schools perform well across tests suggesting there is a real relationship between learning and education. One interpretation of this is that some students learn more (and some schools teach more) of the kinds of things that show up across tests. Another interpretation is that tests are biased. The students and schools that do well also tend to be wealthier and whiter communities which are also where test-writers come from. This

shared cultural capital means that test-makers and students share social and culture references enabling them to interpret a passage about lacrosse or a math question involving various amounts of quinoa without getting confused as to what these are. A final interpretation is that these schools and students simply have more resources to learn with. They spend more time in learning-rich activities like reading, formal instruction, and guided play. These test scores don't reflect shared cultural capital or mastery of some core content, but that they learn more in total.

To return to the NAEP example, what is more likely the case is that we could come up with one or two tests that show American students trouncing their international peers or learning ten times as much now as a generation ago. But 9 out of 10 tests would show the same findings we live with today. The question is whether the one test assesses the kind of education we care about or how it is that we still underperform across all nine tests.

The problem with testing is not that it is arbitrary, but that it doesn't help us answer these basic questions. Right now, we track very little in the way of education. The vast majority of students in the U.S. are given standardized tests in Math and English, and sometimes science every 3-4 years. NAEP is administered to a nationally representative sample of students in three grades every 2-5 years. There is no useful data on what students are learning for the vast majority of subjects, teachers, or students for most years. This is generally fine for the teacher and the student, but severely limits what we can know about what works and what doesn't.

In sum, so long as tests change and schools react; standardized test scores will tell us very little about what works in education. If we want to know whether a school, teacher, or policy is effective at improving education, we must stabilize our tests and decouple them from accountability. We also cannot reify one single test, whether a state's standardized test, PISA and NAEP, or the SAT as the baseline measure of true learning. None is a true measure of education

and each has made its own social and political choices about what to assess and how. To compensate, we must measure learning more ubiquitously and more frequently – in every class, every quarter, for every grade. This call for standardized testing everywhere might seem sadistic, but teachers at Achievement gave quizzes every week and tests every month or two. I'm not suggesting comprehensive exams every quarter, rather we should look to standardize teachers' quizzes and exams. Testing is already a ubiquitous part of the contemporary classroom. It is simply not nearly as useful as it could be in helping us measure and improve education for everyone. Only with more and better data about learning can we begin to get a grip on what actually works.

BUILDING EFFECTIVE TEACHERS

Achievement College Prep worked hard to find, train, and promote the best teachers. They screened them for the right beliefs and practices, devoted substantial time and energy to monitoring, coaching, and supporting their performance in class, and reviewed each teacher's performance every year to decide whether they should stay on. These high expectations and loads of responsibility created a pressure cooker, driving teachers to work long hours, become disillusioned and eventually burn out. During my focal year, they were trying to improve work life balance and hire teachers who weren't all twenty-five-year-old white women teaching for the first time. Yet, teachers still burned out. They did manage to hire a broader set of teachers, but they did not change the fundamental expectations for who those teachers should be.

At the heart of this churn and burn system is the school's idea of the right teacher as a mission-driven, always-improving perfectionist. This ideal influenced the way in which they hired, created their professional community, and decided who should stay and how much they

would be paid. They built a school for this kind of teacher and the ideal continued to replicate itself despite active attempts by school leaders to create work/life balance. Looked at in isolation, these all seemed like good ideas – hire people who are already committed to the organization's mission, provide professional development to help them to get better and achieve that mission, and cut loose those don't perform. Taken together however, they created a burnout factory.

Lesson 4: Good ideas do not necessarily play well together.

The combination of hiring true-believer staff, embedding them in coaching programs, and retaining those who prove to be effective all sounds like a good idea. But taken together, they drive teachers to overwork. As researchers and policymakers, we treat policy ideas as single instruments, a later school start time, higher salaries, teacher coaching. We rarely consider how the whole fits together. Professional community is a great thing to foster in a school, but because teachers are always being observed and developed, it can be difficult to separate from work. For example, teachers had to have their week's lesson plan material ready ahead of time for their weekly review meetings with chairs. While this set a good expectation, it gave teachers little slack if they had to re-engineer their lesson or got sick.

Combine this with teachers who believe in the mission – who feel compelled to put together quality lesson plans and receive feedback as part of their providing excellent education to students who are already behind. These teachers will be much less willing to take a break or write off a coaching session because something came up or didn't work out. Instead, they work nights and weekends to catch up and, more generally, put together excellent lessons.

Finally, when the cost of not performing is that you lose your job, then overwork is seen as the precondition for being an excellent teacher. Inadequacy is the feeling true believers get

when they aren't working 80 hours a week. And, in the absence of a raise, the message is clear – intrinsic motivation is what's most important.

Researchers, policymakers, and school leaders need to understand the knock-on effects our ideas have. For example, professional community likely helps new teachers learn the ropes quickly and maintain well-organized classrooms. But, to prevent burnout, professional community must allow teachers to set boundaries with one another and the school. Part of the community must be the norm that work ends and people go home. And, if the professional community is tied to hiring based on who's a true believer, then teachers must see work moderation as part of being a great teacher and coaches must accept that teachers can only do so much in a day.

Lesson 5: New organizations present unique challenges for embedding best practices

One of the reasons the pay-for-performance system at Achievement led to burnout was that it was poorly planned out. The original idea was that all teachers at ACP, because they were asked back, were "effective" and therefore should be paid the same. When the school decided to pay more for teachers in their third, fourth, and fifth year; they did it as a bonus rather than an increase in pay. This odd initial approach to pay and incremental band-aids are typical of all organizations. However, new organizations suffer from them to an extreme degree because, being new, much of the organization's policies and practices are done this way.

At its founding, Achievement based its competitive advantage on a no-excuses model of excellence which matched intensive instruction to high expectations. No-excuses schools did not come with a model of human resources. Nor did it come with a model of pedagogy, fundraising, student recruitment. These essential functions of the school were built from scratch on an ad hoc

basis. Frankly its surprising so many new schools have survived for so long (the 5-year survival rate for local businesses is typically below 10%). That there is a strong culture, working professional community, and school-wide shift to Common Core are all amazing feats for such a young organization. The poor work/life boundaries, pay scheme, and hiring practices that turn professional community into a pressure cooker were ad hoc decisions the school had to make to fill its ranks.

Researchers, policymakers, and school leaders need to be wary of how best practices work in new schools. Some may be too infeasible for young organizations or too dependent on other school programs and initiatives which new schools may not have or may not yet do well. Moreover, new schools need more support from the state and more regulation to ensure that these ad hoc decisions do not undermine the school itself. If we're going to keep experimenting with new school organizations, we must provide more administrative supports for charter schools and more guard rails to ensure the right things get done for organizations. Moreover, innovation requires the room to experiment, but the experiment needs to be done within principled and well-understood boundaries. If no-excuses schools fail because they never got their hiring right, then what have we learned about holding students to high expectations?

BUILDING EFFECTIVE SCHOOLS

Where do we go from here? Our science isn't bad, our program design and validation is. We don't develop programs that work in schools. We don't tie accountability to real actions schools can take that are tied to meaningful measures. We don't examine our ideas in the contexts in which they're being used. We treat the world as a linear model – put the variable in,

and you get your outcome. Implementation means organizational and institutional change and we have no account for this.

But we shouldn't be surprised by this. Applied research has always been seen as second-rate science in the social sciences. Science does not soil itself by becoming engineering. It is always engaging in some degree of construction in the world. However, as this dissertation has shown, this construction process is much more important than we've given credit. The empirical standards for translating research into practice are much higher than for pure science.

Let's look at an example of success. In 1973, Fischer Black, Myron Scholes, and Robert Merton developed mathematical models for predicting the price of derivatives (Black and Scholes 1973; Merton 1973). When they first developed their model, market prices looked nothing like what they predicted. The saw an opportunity to beat the market and created their own company to trade derivatives and sell their predicted prices to other traders. Other traders began using these prices. Markets themselves built institutions like loans for derivative trading and policies like short selling which made the market more like the abstract market Black and Scholes described a decade earlier. Over time, the price of derivatives began to look more and more like Black, Scholes, and Merton's predictions (MacKenzie and Millo 2003). In 1995, they won the Nobel Prize for their work and their work is considered an essential part of the canon of economic knowledge.

Scholars who study science have long critiqued the veil of objectivity projected by scientists. People like Bruno Latour (1987) and Michel Callon (2007) have gone so far as to say that scientists create the world around them. Black, Scholes, and Merton could be seen as cheating at the truth. Rather than know something and have that knowledge proven true, they are only right because they went out and reshaped the world to fit their theory.

Latour, Callon, and others are not criticizing the validity of the science or calling them epistemic cheaters. They are recognizing the continuity between lab and life. They are criticizing the myth that science observes the world from an objective distance and is good to the extent that it reflects the world in its pristine form. In every article, grant proposal, or job application; researchers are asked "who cares" and "so what." These questions, though ubiquitous and fundamental to the enterprise, are almost always answered in the hypothetical. "If schools reduce class sizes by ten percent, students test scores would go up 5%." We care about impact but do little to develop it and even less to reward it. In organization theory, we'd say our concerns with impact are decoupled from how we actually conduct our research (Meyer and Rowan 1977). We ceremoniously answer these questions about impact with the myths we tell ourselves about how research translates into practice. This decoupling insulates us from having to actually prove our knowledge is either true or impactful.

The reality is that most of our research is likely not impactful. That's certainly the track record in most fields with strong histories of success. As Thomas Edison said "I have constructed three thousand different theories in connection with the electric light, each one of them reasonable and apparently to be true. Yet only in two cases did my experiments prove the truth of my theory." This rate of success conforms to present-day trials in mice (Seok et al. 2013; van der Worp et al. 2010). We should assume the social sciences are no different and take it with a grain of salt when someone claims to find some variable that matters.

We need to redefine social science not to be the production of knowledge, but as the production of *useful* knowledge. Identifying some mechanism, pattern, or idea within in observational data is only the first stage of the new Science. In order to determine whether this insight really matters, we need engineer solutions that use it. We should apply the same level of

rigor to assessing the implementation of our solution to the problem of the subject as we did the original result. We should expect many of these implementations to fail. If enough fail, perhaps the theory was not as insightful as we thought. Or, as some succeed, we look to generalize the application to more schools. We scale up.

To the extent we see science as the pursuit of objective knowledge about a pristine world, we give up our ability to shape the world. Science is a co-creator of the world. To return to the example of derivatives, economists did not create derivatives markets. They sold their model to existing traders and eventually created their own company. Their model became the sine qua non for predicting, and therefore setting, derivatives prices. They also advocated for new rules and legislation to make markets a bit more like their theories. As a result, they regularized prices and created liquidity, stabilizing a market that is now a cornerstone for mitigating risk. They were not right about the market when they created their ideas, but they created a market that eventually converged with their theory and solved real social problems. Derivatives markets are certainly not perfect. But they work well most of the time. Shouldn't all our science be so effective?

BIBLIOGRAPHY

Aaronson, Daniel, Lisa Barrow, and William Sander. 2007. "Teachers and Student Achievement in the Chicago Public High Schools." *Journal of Labor Economics* 25 (1): 95–135. https://doi.org/10.1086/508733.

Abbott, Andrew. 1988a. *The System of Professions*. Chicago: University of Chicago Press.

———. 1988b. "Transcending General Linear Reality." *Sociological Theory* 6 (2): 169. https://doi.org/10.2307/202114.

Aladjem, Daniel K, Beatrice F Birman, Martin Orland, Jenifer Harr-Robins, Alberto Heredia, Thomas B Parrish, and Stephen J Ruffini. 2010. "Achieving Dramatic School Improvement: An Exploratory Study. A Cross-Site Analysis from the Evaluation of Comprehensive School Reform Program Implementation and Outcomes Study." Washington, D.C.: U.S. Department of Education.

Aladjem, Daniel K, Kerstin Carlson LeFloch, Yu Zhang, Anja Kurki, Andrea Boyle, James E Taylor, Suzannah Herrmann, Kazuaki Uekawa, Kerri Thomsen, and Olatokunbo Fashola. 2006. "Models Matter–The Final Report of the National Longitudinal Evaluation of Comprehensive School Reform." Washington, D.C.: American Institutes for Research. http://files.eric.ed.gov/fulltext/ED499198.pdf.

Amrein-Beardsley, Audrey, and David C. Berliner. 2002. "The Impact of High-Stakes Tests on Student Academic Performance." http://nepc.colorado.edu/files/EPSL-0211-126-EPRU.pdf.

Ancona, Deborah, and Chee-Leong Chong. 1996. "Entrainment: Pace, Cycle, and Rhythm in Organizational Behavior." *Research in Organizational Behavior* 18: 251–84.

Andersen, Torben Juul, and Bo Bernhard Nielsen. 2009. "Adaptive Strategy Making: The Effects of Emergent and Intended Strategy Modes." *European Management Review* 6 (2): 94–106.

Anderson, Philip. 1999. "Perspective: Complexity Theory and Organization Science." *Organization Science* 10 (3): 216–32. https://doi.org/10.1287/orsc.10.3.216.

Angrist, Joshua D. 2004. "American Education Research Changes Tack." *Oxford Review of Economic Policy* 20 (2): 198–212. https://doi.org/10.1093/oxrep/grh011.

Angrist, Joshua D., Eric Bettinger, and Michael Kremer. 2006. "Long-Term Educational Consequences of Secondary School Vouchers: Evidence from Administrative Records in Colombia." *American Economic Review* 96 (3): 847–62. https://doi.org/10.1257/aer.96.3.847.

Angrist, Joshua D., Peter D. Hull, Parag A. Pathak, and Christopher R. Walters. 2017. "Leveraging Lotteries for School Value-Added: Testing and Estimation*." *The Quarterly Journal of Economics* 132 (2): 871–919. https://doi.org/10.1093/qje/qjx001.

Angrist, Joshua D., Parag A. Pathak, and Christopher R. Walters. 2013. "Explaining Charter School Effectiveness." *American Economic Journal: Applied Economics* 5 (4): 1–27.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2010. "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics." *Journal of Economic Perspectives* 24 (2): 3–30. https://doi.org/10.1257/jep.24.2.3.

Arsen, David, and Yongmei Ni. 2012. "The Effects of Charter School Competition on School District Resource Allocation." *Educational Administration Quarterly* 48 (1): 3–38. https://doi.org/10.1177/0013161X11419654.

Bartley, Tim, and Niklas Egels-Zandén. 2016. "Beyond Decoupling: Unions and the Leveraging of Corporate Social Responsibility in Indonesia." *Socio-Economic Review* 14 (2): 231–55. https://doi.org/10.1093/ser/mwv023.

Barzelay, Michael. 1992. *Breaking through Bureaucracy: A New Vision for Managing in Government*. Univ of California Press.

Battilana, Julie, Bernard Leca, and Eva Boxenbaum. 2009. "How Actors Change Institutions: Towards a Theory of Institutional Entrepreneurship." *Academy of Management Annals* 3 (1): 65–107. https://doi.org/10.1080/19416520903053598.

Berends, Mark. 2015. "Sociology and School Choice: What We Know After Two Decades of Charter Schools." *Annual Review of Sociology* 41 (1): 159–80. https://doi.org/10.1146/annurev-soc-073014-112340.

Berends, Mark, Ellen Goldring, Marc Stein, and Xiu Cravens. 2010. "Instructional Conditions in Charter Schools and Students' Mathematics Achievement Gains." *American Journal of Education* 116 (3): 303–35. https://doi.org/10.1086/651411.

Betts, Julian R, and Y Emily Tang. 2019. "The Effect of Charter Schools on Student Achievement: A Meta-Analysis of the Literature." In *School Choice at the Crossroads: Research Perspectives*, edited by Mark Berends, R. Joseph Waddington, and John Shoenig, 67–89. New York: Routledge.

Betts, Julian R, Andrew Zau, and Lorien Rice. 2003. "Determinants of Student Achievement: New Evidence from San Diego." Public Policy Institute of California.

Bigley, Gregory A., and Karlene H. Roberts. 2001. "The Incident Command System: High-Reliability Organizing for Complex and Volatile Task Environments." *Academy of Management Journal* 44 (6): 1281–99.

Black, Fischer, and Myron Scholes. 1973. “The Pricing of Options and Corporate Liabilities.” *Journal of Political Economy* 81 (3): 637–54. https://doi.org/10.1086/260062.

Borman, Geoffrey D, Gina M Hewes, Laura T Overman, and Shelly Brown. 2003. “Comprehensive School Reform and Achievement: A Meta-Analysis.” *Review of Educational Research* 73 (2): 125–230.

Bourdieu, Pierre. 1977. *Outline of a Theory of Practice*. Translated by Richard Nice. Cambridge, UK: Cambridge University Press.

Bowles, Samuel, and Herbert Gintis. 1976. *Schooling in Capitalist America*. New York: Basic Books.

Bromley, Patricia, Hokyu Hwang, and Walter W. Powell. 2012. “Decoupling revisited: Common pressures, divergent strategies in the U. S. nonprofit sector.” *M@n@gement* 15 (5): 469. https://doi.org/10.3917/mana.155.0469.

Bromley, Patricia, and Walter W. Powell. 2012. “From Smoke and Mirrors to Walking the Talk: Decoupling in the Contemporary World.” *The Academy of Management Annals* 6 (1): 483–530. https://doi.org/10.1080/19416520.2012.684462.

Bryk, Anthony, Eric Camburn, and Karen Seashore Louis. 1999. “Professional Community in Chicago Elementary Schools: Facilitating Factors and Organizational Consequences.” *Educational Administration Quarterly* 35 (5): 751–781.

Bryk, Anthony S., and Barbara Schneider. 2003. “Trust in Schools: A Core Resource for School Reform.” *Educational Leadership* 60 (6): 40–45.

Burgelman, Robert A. 1983. “A Model of the Interaction of Strategic Behavior, Corporate Context, and the Concept of Strategy.” *The Academy of Management Review* 8 (1): 61–70.

Butler, Judith. 1988. “Performative Acts and Gender Constitution: An Essay in Phenomenology and Feminist Theory.” *Theatre Journal* 40 (4): 519–531.

Cabantous, Laure, and Jean-Pascal Gond. 2011. “Rational Decision Making as Performative Praxis: Explaining Rationality’s Éternel Retour.” *Organization Science* 22 (3): 573–86. https://doi.org/10.1287/orsc.1100.0534.

Callon, Michel. 2007. “What Does It Mean to Say That Economics Is Performative?” In *Do Economists Make Markets? On the Performance of Economics*, edited by Donald MacKenzie, Fabian Muniesa, and Lucian Siu, 311–58. Princeton, NJ: Princeton University Press.

Campo, Charlotte a. 1993. “Collaborative School Cultures: How Principals Make A Difference.” *School Leadership & Management* 13 (2): 119–27. https://doi.org/10.1080/0260136930130202.

Carnoy, Martin, Richard F Elmore, and Leslie Santee Siskin. 2003. *The New Accountability: High Schools and High Stakes Testing*. New York: RoutledgeFalmer. http://public.eblib.com/choice/publicfullrecord.aspx?p=198528.

Carnoy, Martin, and Susanna Loeb. 2002. "Does External Accountability Affect Student Outcomes? A Cross-State Analysis." *Educational Evaluation and Policy Analysis* 24 (4): 305–331.

Carter, Samuel Casey. 2000. *No Excuses: Lessons from 21 High-Performing, High-Poverty Schools*. Washington, D.C.: Heritage Foundation.

Center for Research on Education Outcomes (CREDO). 2009. "Multiple Choice: Charter School Performance in 16 States." Stanford, CA: CREDO.

Chatman, Jennifer A., and Charles A. O'Reilly. 2016. "Paradigm Lost: Reinvigorating the Study of Organizational Culture." *Research in Organizational Behavior* 36: 199–224. https://doi.org/10.1016/j.riob.2016.11.004.

Chen, K. K. 2013. "Storytelling: An Informal Mechanism of Accountability for Voluntary Organizations." *Nonprofit and Voluntary Sector Quarterly* 42 (5): 902–22. https://doi.org/10.1177/0899764012455699.

Cheng, Albert, Collin Hitt, Brian Kisida, and Jonathan N. Mills. 2017. "'No Excuses' Charter Schools: A Meta-Analysis of the Experimental Evidence on Student Achievement." *Journal of School Choice* 11 (2): 209–38. https://doi.org/10.1080/15582159.2017.1286210.

Cherniss, Cary. 1980. *Professional Burnout in Human Service Organizations*. New York: Praeger.

Child, John. 1997. "Strategic Choice in the Analysis of Action, Structure, Organizations and Environment: Retrospect and Prospect." *Organization Studies* 18 (1): 43–76.

Clark, Melissa A., Philip M. Gleason, Christina Clark Tuttle, and Marsha K. Silverberg. 2015. "Do Charter Schools Improve Student Achievement?" *Educational Evaluation and Policy Analysis* 37 (4): 419–36. https://doi.org/10.3102/0162373714558292.

Clayton, Martin. 2012. "What Is Entrainment? Definition and Applications in Musical Research." *Empirical Musicology Review* 7 (1–2): 49–56.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2006. "Teacher-Student Matching and the Assessment of Teacher Effectiveness." *Journal of Human Resources* XLI (4): 778–820. https://doi.org/10.3368/jhr.XLI.4.778.

———. 2010. "Teacher Credentials and Student Achievement in High School: A Cross-Subject Analysis with Student Fixed Effects." *Journal of Human Resources* 45 (3): 655–81. https://doi.org/10.3368/jhr.45.3.655.

Coburn, Cynthia E. 2004. “Beyond Decoupling: Rethinking the Relationship Between the Institutional Environment and the Classroom.” *Sociology of Education* 77 (3): 211–44. https://doi.org/10.1177/003804070407700302.

Cohen, David K. 1990. “A Revolution in One Classroom: The Case of Mrs. Oublier.” *Educational Evaluation and Policy Analysis* 12 (3): 311–29. https://doi.org/10.3102/01623737012003311.

Cohen, Michael D., James G. March, and Johan P. Olsen. 1972. “A Garbage Can Model of Organizational Choice.” *Administrative Science Quarterly* 17 (1): 1–25.

Coleman, James S, Ernest Q Campbell, Carol J Hobson, James McPartland, Alexander M Mood, Frederic D Weinfeld, and Robert York. 1966. *Equality of Educational Opportunity*. Washington: U.S. Department of Health, Education, and Welfare.

Coleman, James S, and John Wallace Claire Johnstone. 1963. *The Adolescent Society*. Oxford: Free Press of Glenco.

Coleman, James Samuel, Thomas Hoffer, and Sally Kilgore. 1982. *High School Achievement: Public, Catholic, and Private Schools Compared*. Basic Books New York.

Cuban, Larry, and David Tyack. 1995. *Tinkering toward Utopia: A Century of Public School Reform*. Cambridge, MA: Harvard University Press.

Cullen, Julie Berry, Brian A Jacob, and Steven Levitt. 2006. “The Effect of School Choice on Participants: Evidence from Randomized Lotteries.” *Econometrica* 74 (5): 1191–1230. https://doi.org/10.1111/j.1468-0262.2006.00702.x.

Cyert, Richard Michael, and James G. March. 1992. *A Behavioral Theory of the Firm*. 2nd ed. Cambridge, Mass., USA: Blackwell Business.

D’Adderio, Luciana, and Neil Pollock. 2014. “Performing Modularity: Competing Rules, Performative Struggles and the Effect of Organizational Theories on the Organization.” *Organization Studies* 35 (12): 1813–1843.

Davis, Dave, Mary Ann Thomson O’Brien, Nick Freemantle, Fredric Wolf, Paul Mazmanian, and Anne Taylor-Vaisey. 1999. “Impact of Formal Continuing Medical Education: Do Conferences, Workshops, Rounds, and Other Traditional Continuing Education Activities Change Physician Behavior or Health Care Outcomes?” *JAMA* 282 (9): 867–74. https://doi.org/10.1001/jama.282.9.867.

Davis, Tomeka M. 2013. “Charter School Competition, Organization, and Achievement in Traditional Public Schools.” *Education Policy Analysis Archives* 21 (December): 88. https://doi.org/10.14507/epaa.v21n88.2013.

Deal, Terrence E., and Lynn D. Celotti. 1980. “How Much Influence Do (And Can) Educational Administrators Have on Classrooms?” *The Phi Delta Kappan* 61 (7): 471–473.

Deal, Terrence E, and Allan A Kennedy. 1983. “Culture and School Performance.” *Educational Leadership* 40 (5): 14–15.

Deal, Terrence E, and Kent D Peterson. 1999. *Shaping School Culture: The Heart of Leadership.* San Francisco, CA: Jossey-Bass.

Dee, Thomas S., and Brian Jacob. 2011. “The Impact of No Child Left Behind on Student Achievement.” *Journal of Policy Analysis and Management* 30 (3): 418–446. https://doi.org/10.1002/pam.20586.

Dee, Thomas S., and Benjamin J. Keys. 2004. “Does Merit Pay Reward Good Teachers? Evidence from a Randomized Experiment.” *Journal of Policy Analysis and Management* 23 (3): 471–488. https://doi.org/10.1002/pam.20022.

Diamond, John B. 2007. “Where the Rubber Meets the Road: Rethinking the Connection Between High-Stakes Testing Policy and Classroom Instruction.” *Sociology of Education* 80 (4): 285–313. https://doi.org/10.1177/003804070708000401.

DiMaggio, Paul J, and Walter W Powell. 1983. “The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields.” *American Sociological Review* 48 (2): 147–160.

Dobbie, Will, and Roland G Fryer. 2013. “Getting beneath the Veil of Effective Schools: Evidence from New York City.” *American Economic Journal: Applied Economics* 5 (4): 28–60.

Dobbin, Frank, and Jiwook Jung. 2010. “The Misapplication of Mr. Michael Jensen: How Agency Theory Brought down the Economy and Why It Might Again.” *Research in the Sociology of Organizations* 30: 29–64.

Dobbin, Frank, Soohan Kim, and Alexandra Kalev. 2011. “You Can’t Always Get What You Need: Organizational Determinants of Diversity Programs.” *American Sociological Review* 76 (3): 386–411. https://doi.org/10.1177/0003122411409704.

Dobbin, Frank, and Dirk Zorn. 2005. “Corporate Malfeasance and the Myth of Shareholder Value.” *Political Power and Social Theory* 17: 179–198.

Douglas, Heather E. 2009. “Reintroducing Prediction to Explanation.” *Philosophy of Science* 76 (4): 444–63. https://doi.org/10.1086/648111.

Dunleavy, P., Helen Margetts, Simon Bastow, and Jane Tinkler. 2005. “New Public Management Is Dead--Long Live Digital-Era Governance.” *Journal of Public Administration Research and Theory* 16 (3): 467–94. https://doi.org/10.1093/jopart/mui057.

Eisenhardt, Kathleen M. 1989. “Building Theories from Case Study Research.” *Academy of Management Review* 14 (4): 532–50. https://doi.org/10.5465/AMR.1989.4308385.

Eisenhardt, Kathleen M., and Melissa E. Graebner. 2007. “Theory Building from Cases: Opportunities and Challenges.” *Academy of Management Journal* 50 (1): 25–32.

Fama, Eugene F. 1980. “Agency Problems and the Theory of the Firm.” *The Journal of Political Economy*, 288–307.

Fama, Eugene F., and Michael C. Jensen. 1983. “Separation of Ownership and Control.” *The Journal of Law & Economics* 26 (2): 301–25.

Feldman, Daniel C., and Thomas W. H. Ng. 2007. “Careers: Mobility, Embeddedness, and Success.” *Journal of Management* 33 (3): 350–77. https://doi.org/10.1177/0149206307300815.

Feldman, Martha S, and Brian T Pentland. 2003. “Reconceptualizing Organizational Routines as a Source of Flexibility and Change.” *Administrative Science Quarterly* 48 (1): 94–118.

Ferraro, Fabrizio, Jeffrey Pfeffer, and Robert I. Sutton. 2005. “Economics Language and Assumptions: How Theories Can Become Self-Fulfilling.” *Academy of Management Review* 30 (1): 8–24.

———. 2009. “How and Why Theories Matter: A Comment on Felin and Foss (2009).” *Organization Science* 20 (3): 669–75. https://doi.org/10.1287/orsc.1090.0432.

Finnigan, Kara S. 2007. “Charter School Autonomy: The Mismatch Between Theory and Practice.” *Educational Policy* 21 (3): 503–26. https://doi.org/10.1177/0895904806289189.

Firestone, William A, and Bruce L Wilson. 1985. “Using Bureaucratic and Cultural Linkages to Improve Instruction: The Principal’s Contribution.” *Educational Administration Quarterly* 21 (2): 7–30.

Freeman, Scott, Sarah L. Eddy, Miles McDonough, Michelle K. Smith, Nnadozie Okoroafor, Hannah Jordt, and Mary Pat Wenderoth. 2014. “Active Learning Increases Student Performance in Science, Engineering, and Mathematics.” *Proceedings of the National Academy of Sciences* 111 (23): 8410–15. https://doi.org/10.1073/pnas.1319030111.

Freudenberger, Herbert J. 1975. “The Staff Burn-out Syndrome in Alternative Institutions.” *Psychotherapy: Theory, Research & Practice* 12 (1): 73–82. https://doi.org/10.1037/h0086411.

Furgeson, Joshua, Brian Gill, Joshua Haimson, Alexandra Killewald, Moira McCullough, Ira Nichols-Barrer, Natalya Verbitsky-Savitz, et al. 2012. “Charter-School Management Organizations: Diverse Strategies and Diverse Student Impacts.” *Mathematica Policy Research, Inc.* http://eric.ed.gov/?id=ED528536.

Fusaroli, Riccardo, Marcus Perlman, Alan Mislove, Alexandra Paxton, Teenie Matlock, and Rick Dale. 2015. "Timescales of Massive Human Entrainment." *PLOS ONE* 10 (4): 1–19. https://doi.org/10.1371/journal.pone.0122742.

Gehman, J., L. K. Trevino, and R. Garud. 2013. "Values Work: A Process Study of the Emergence and Performance of Organizational Values Practices." *Academy of Management Journal* 56 (1): 84–112. https://doi.org/10.5465/amj.2010.0628.

Gehman, Joel, and Jean-François Soublière. 2017. "Cultural Entrepreneurship: From Making Culture to Cultural Making." *Innovation* 19 (1): 61–73. https://doi.org/10.1080/14479338.2016.1268521.

Gersten, Russell, Mary Jo Taylor, Tran D Keys, Eric Rolfhus, and Rebecca Newman-Gonchar. 2014. "Summary of Research on the Effectiveness of Math Professional Development Approaches. REL 2014-010." Washington, D.C.: Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast.

Glaser, Barney G., and Anselm L. Strauss. 2017. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. 1st ed. Routledge. https://doi.org/10.4324/9780203793206.

Glaser, Vern L., Peer C. Fiss, and Mark Thomas Kennedy. 2016. "Making Snowflakes Like Stocks: Stretching, Bending, and Positioning to Make Financial Market Analogies Work in Online Advertising." *Organization Science* 27 (4): 1029–48. https://doi.org/10.1287/orsc.2016.1069.

Gleason, Philip, Melissa Clark, Christina Clark Tuttle, and Emily Dwoyer. 2010. "The Evaluation of Charter School Impacts." Mathematica Policy Research.

Golann, Joanne W. 2015. "The Paradox of Success at a No-Excuses School." *Sociology of Education* 88 (2): 103–19. https://doi.org/10.1177/0038040714567866.

Goldring, Rebecca, Soheyla Taie, and Minsun Riddles. 2014. "Teacher Attrition and Mobility: Results from the 2012-13 Teacher Follow-Up Survey. First Look. NCES 2014-077." *National Center for Education Statistics*. http://eric.ed.gov/?id=ED546773.

Goldstein, Jeffrey. 1999. "Emergence as a Construct: History and Issues." *Emergence* 1 (1): 49–72. https://doi.org/10.1207/s15327000em0101_4.

Gond, Jean-Pascal, Laure Cabantous, Nancy Harding, and Mark Learmonth. 2016. "What Do We Mean by Performativity in Organizational and Management Theory? The Uses and Abuses of Performativity: Organizing Performativity." *International Journal of Management Reviews* 18 (4): 440–63. https://doi.org/10.1111/ijmr.12074.

Goodlad, John I. 1984. *A Place Called School: Prospects for the Future.* New York: McGraw-Hill.

Goodman, J. F. 2013. “Charter Management Organizations and the Regulated Environment: Is It Worth the Price?” *Educational Researcher* 42 (2): 89–96. https://doi.org/10.3102/0013189X12470856.

Graham, John R., Campbell R. Harvey, Jillian A. Popadak, and Shivaram Rajgopal. 2016. “Corporate Culture: Evidence from the Field.” *SSRN*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2805602.

Grodsky, Eric, John Robert Warren, and Erika Felts. 2008. “Testing and Social Stratification in American Education.” *Annual Review of Sociology* 34 (1): 385–404. https://doi.org/10.1146/annurev.soc.34.040507.134711.

Haack, Patrick, Dennis Schoeneborn, and Christopher Wickert. 2012. “Talking the Talk, Moral Entrapment, Creeping Commitment? Exploring Narrative Dynamics in Corporate Responsibility Standardization.” *Organization Studies* 33 (5–6): 815–45. https://doi.org/10.1177/0170840612443630.

Hafner-Burton, Emilie M., and Kiyoteru Tsutsui. 2005. “Human Rights in a Globalizing World: The Paradox of Empty Promises.” *American Journal of Sociology* 110 (5): 1373–1411. https://doi.org/10.1086/428442.

Hallett, Tim, and Marc J. Ventresca. 2006. “Inhabited Institutions: Social Interactions and Organizational Forms in Gouldner’s Patterns of Industrial Bureaucracy.” *Theory and Society* 35 (2): 213–36. https://doi.org/10.1007/s11186-006-9003-z.

Hallett, Timothy. 2010. “The Myth Incarnate: Recoupling Processes, Turmoil, and Inhabited Institutions in an Urban Elementary School.” *American Sociological Review* 75 (1): 52–74. https://doi.org/10.1177/0003122409357044.

Hamilton, Laura, Brian Stecher, and Kun Yuan. 2008. “Standards-Based Reform in the United States: History, Research, and Future Directions.” Santa Monica, CA: Rand Education.

Hannan, Michael T, and John Freeman. 1977. “The Population Ecology of Organizations.” *American Journal of Sociology* 82 (5): 929–964.

Hannan, Michael T., and John Freeman. 1984. “Structural Inertia and Organizational Change.” *American Sociological Review* 49 (2): 149–64.

Hanushek, Eric A., and Margaret E. Raymond. 2005. “Does School Accountability Lead to Improved Student Performance?” *Journal of Policy Analysis and Management* 24 (2): 297–327. https://doi.org/10.1002/pam.20091.

Harris, Douglas N., and Tim R. Sass. 2011. “Teacher Training, Teacher Quality and Student Achievement.” *Journal of Public Economics* 95 (7): 798–812.

Harrison, J Richard, and Glenn Carroll. 2006. *Culture and Demography in Organizations*. Princeton, NJ: Princeton University Press.

Haveman, Heather A. 1992. “Between a Rock and a Hard Place: Organizational Change and Performance Under Conditions of Fundamental Environmental Transformation.” *Administrative Science Quarterly* 37 (1): 48. https://doi.org/10.2307/2393533.

———. 1993. “Follow the Leader: Mimetic Isomorphism and Entry Into New Markets.” *Administrative Science Quarterly* 38 (4): 593–627.

Herman, Rebecca. 1999. “An Educators’ Guide to School Reform.” Washington, DC: American Institutes for Research.

Hill, Heather C., Mary Beisiegel, and Robin Jacob. 2013. “Professional Development Research: Consensus, Crossroads, and Challenges.” *Educational Researcher* 42 (9): 476–487.

Hirschfield, P. J. 2008. “Preparing for Prison?: The Criminalization of School Discipline in the USA.” *Theoretical Criminology* 12 (1): 79–101. https://doi.org/10.1177/1362480607085795.

Hochschild, Arlie Russell. 1983. *The Managed Heart: Commercialization of Human Feeling*. Berkeley: Univ of California Press.

Hoxby, Caroline M. 2002. “Would School Choice Change the Teaching Profession?” *The Journal of Human Resources* 37 (4): 846. https://doi.org/10.2307/3069619.

Hoxby, Caroline M., Sonali Murarka, and Jenny Kang. 2009. “How New York City’s Charter Schools Affect Achievement.” *Cambridge, MA: New York City Charter Schools Evaluation Project*. http://www.nber.org/~schools/charterschoolseval/how_NYC_charter_schools_affect_achievement_sept2009.pdf.

Ingersoll, Richard M. 2004. “Why Do High-Poverty Schools Have Difficulty Staffing Their Classrooms with Qualified Teachers?” Center for American Progress, Institute for America’s Future. https://scholar.gse.upenn.edu/rmi/files/ingersoll-final.pdf.

Ingersoll, Richard, Lisa Merrill, and Daniel Stuckey. 2014. “Seven Trends: The Transformation of the Teaching Force.” http://repository.upenn.edu/cpre_researchreports/79/.

Jackson, C. Kirabo, Rucker C. Johnson, and Claudia Persico. 2016. “The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms.” *The Quarterly Journal of Economics* 131 (1): 157–218. https://doi.org/10.1093/qje/qjv036.

Jackson, Michelle, and D.R. Cox. 2013. “The Principles of Experimental Design and Their Application in Sociology.” *Annual Review of Sociology* 39 (1): 27–49. https://doi.org/10.1146/annurev-soc-071811-145443.

James, Oliver, and Nick Manning. 1996. “Public Management Reform: A Global Perspective.” *Politics* 16 (3): 143–49. https://doi.org/10.1111/j.1467-9256.1996.tb00033.x.

Jennings, Jennifer L., and Jonathan Marc Bearak. 2014. "'Teaching to the Test' in the NCLB Era." *Educational Researcher* 43 (8): 381–89. https://doi.org/10.3102/0013189X14554449.

Jennings, Jennifer L, and Douglas Lee Lauen. 2016. "Accountability, Inequality, and Achievement: The Effects of the No Child Left Behind Act on Multiple Measures of Student Learning." *RSF: The Russell Sage Foundation Journal of the Social Sciences* 2 (5): 220–241.

Jensen, Michael C, and William H Meckling. 1976. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *Journal of Financial Economics* 3 (4): 305–360.

Jepperson, Ronald L. 1991. "Institutions, Institutional Effects, and Institutionalism." In *The New Institutionalism in Organizational Analysis*, edited by Walter W. Powell and Paul J DiMaggio, 143–163. Chicago: University Of Chicago Press.

Jha, Harsh K., and Christine M. Beckman. 2017. "A Patchwork of Identities: Emergence of Charter Schools as a New Organizational Form." In *Research in the Sociology of Organizations*, edited by Marc-David L. Seidel and Henrich R. Greve, 50:69–107. Emerald Publishing Limited. https://doi.org/10.1108/S0733-558X20170000050003.

Jick, Todd D. 1979. "Mixing Qualitative and Quantitative Methods: Triangulation in Action." *Administrative Science Quarterly* 24 (4): 602. https://doi.org/10.2307/2392366.

Kane, Thomas J. 2014. "Do Value-Added Estimates Identify Causal Effects of Teachers and Schools?" The Brown Center Chalkboard. Brookings Institution. https://www.brookings.edu/research/do-value-added-estimates-identify-causal-effects-of-teachers-and-schools/.

———. 2016. "Connecting Practice to Education." *Education Next*, 2016. https://www.educationnext.org/connecting-to-practice-put-education-research-to-work/.

Kane, Thomas J, Jonah E Rockoff, and Douglas O Staiger. 2008. "What Does Certification Tell Us about Teacher Effectiveness? Evidence from New York City." *Economics of Education Review* 27 (6): 615–631.

Keigher, Ashley. 2010. "Teacher Attrition and Mobility: Results from the 2008-09 Teacher Follow-Up Survey. First Look. NCES 2010-353." *National Center for Education Statistics*. http://eric.ed.gov/?id=ED511305.

Kelly, E., and F. Dobbin. 1998. "How Affirmative Action Became Diversity Management: Employer Response to Antidiscrimination Law, 1961-1996." *American Behavioral Scientist* 41: 960–84.

King, Brayden G., Elisabeth S. Clemens, and Melissa Fry. 2011. "Identity Realization and Organizational Forms: Differentiation and Consolidation of Identities among Arizona's Charter Schools." *Organization Science* 22 (3): 554–572.

King, Brayden G., Teppo Felin, and David A. Whetten. 2010. "Perspective—Finding the Organization in Organizational Theory: A Meta-Theory of the Organization as a Social Actor." *Organization Science* 21 (1): 290–305. https://doi.org/10.1287/orsc.1090.0443.

Klein, Stephen, Laura Hamilton, Daniel McCaffrey, and Brian Stecher. 2000. "What Do Test Scores in Texas Tell Us?" *Education Policy Analysis Archives* 8 (0): 49. https://doi.org/10.14507/epaa.v8n49.2000.

Kober, Nancy, Naomi Chudowsky, and Victor Chudowsky. 2008. "Has Student Achievement Increased since 2002? State Test Score Trends through 2006-07." *Center on Education Policy*.

Koretz, Daniel M, and Sheila I Barron. 1998. "The Validity of Gains in Scores on the Kentucky Instructional Results Information System (KIRIS)." Santa Monica, CA: RAND. https://www.rand.org/content/dam/rand/pubs/monograph_reports/2009/MR1014.pdf.

Kornberger, M., and S. Clegg. 2011. "Strategy as Performative Practice: The Case of Sydney 2030." *Strategic Organization* 9 (2): 136–62. https://doi.org/10.1177/1476127011407758.

Kostova, Tatiana, and Kendall Roth. 2002. "Adoption of an Organizational Practice by Subsidiaries of Multinational Corporations: Institutional and Relational Effects." *Academy of Management Journal* 45 (1): 215–233.

Kraaijenbrink, Jeroen, J.-C. Spender, and Aard J. Groen. 2010. "The Resource-Based View: A Review and Assessment of Its Critiques." *Journal of Management* 36 (1): 349–72. https://doi.org/10.1177/0149206309350775.

Kunda, Gideon. 2009. *Engineering Culture: Control and Commitment in a High-Tech Corporation*. Philadelphia, PA: Temple University Press.

Kupchik, Aaron. 2010. *Homeroom Security: School Discipline in an Age of Fear*. New York: New York University Press.

Lake, Robin, Brianna Dusseault, Melissa Bowen, Allison Demeritt, and Paul Hill. 2010. "The National Study of Charter Management Organization (CMO) Effectiveness. Report on Interim Findings." *Center on Reinventing Public Education*. http://eric.ed.gov/?id=ED516865.

Latour, Bruno. 1987. *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge, MA: Harvard university press.

Lawrence, Barbara S. 1997. “Perspective—The Black Box of Organizational Demography.” *Organization Science* 8 (1): 1–22. https://doi.org/10.1287/orsc.8.1.1.

Lee, Jaekyung, and Gary Orfield. 2006. “Tracking Achievement Gaps and Assessing the Impact of NCLB on the Gaps: An In-Depth Look into National and State Reading and Math Outcome Trends.” Cambridge, MA: The Civil Rights Project at Harvard University. http://escholarship.org/uc/item/4db9154t.pdf.

Leiter, Michael P, and Christina Maslach. 2003. “Areas of Worklife: A Structured Approach to Organizational Predictors of Job Burnout.” In *Research in Occupational Stress and Well-Being*, 3:91–134. Bingley: Emerald (MCB UP). https://doi.org/10.1016/S1479-3555(03)03003-8.

Leithwood, Kenneth, and Doris Jantzi. 1990. “Transformational Leadership: How Principals Can Help Reform School Cultures.” *School Effectiveness and School Improvement* 1 (4): 249–80. https://doi.org/10.1080/0924345900010402.

Letiche, Hugo, and Rouven E. Hagemeijer. 2004. “Linkages and Entrainment.” *Journal of Organizational Change Management* 17 (4): 365–82. https://doi.org/10.1108/09534810410545128.

Lewin, Arie Y., Chris P. Long, and Timothy N. Carroll. 1999. “The Coevolution of New Organizational Forms.” *Organization Science* 10 (5): 535–50. https://doi.org/10.1287/orsc.10.5.535.

Lipsky, Michael. 2010. *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services*. 30th anniversary expanded ed. New York: Russell Sage Foundation.

Lortie, Dan C. 1975. *School Teacher: A Sociological Inquiry*. Chicago: University of Chicago Press.

Lounsbury, Michael. 2001. “Institutional Sources of Practice Variation: Staffing College and University Recycling Programs.” *Administrative Science Quarterly* 46 (1): 29–56. https://doi.org/10.2307/2667124.

Lubienski, Christopher. 2003. “Innovation in Education Markets: Theory and Evidence on the Impact of Competition and Choice in Charter Schools.” *American Educational Research Journal* 40 (2): 395–443.

MacKenzie, Donald. 2003. “Long-Term Capital Management and the Sociology of Arbitrage.” *Economy and Society* 32 (3): 349–380.

MacKenzie, Donald, and Yuval Millo. 2003. “Constructing a Market, Performing Theory: The Historical Sociology of a Financial Derivatives Exchange.” *American Journal of Sociology* 109 (1): 107–45.

Malloy, Courtney L, and Priscilla Wohlstetter. 2003. “Working Conditions In Charter Schools What’s the Appeal for Teachers?” *Education and Urban Society* 35 (2): 219–241.

Martin, Joanne. 2002. *Organizational Culture: Mapping the Terrain*. Thousand Oaks, CA: Sage Publications.

Maslach, Christina, Wilmar B. Schaufeli, and Michael P. Leiter. 2001. “Job Burnout.” *Annual Review of Psychology* 52 (1): 397–422. https://doi.org/10.1146/annurev.psych.52.1.397.

McGrath, Joseph E. 1990. “Time Matters in Groups.” In *Intellectual Teamwork: Social and Technological Foundations of Cooperative Work*, edited by J Galegher, R.E. Kraut, and C. Egido, 23:23–61. Hillsdale, NJ: Erblaum. http://web.mit.edu/~yves/www/Missy/TST/90-mcgrath%20-%20time%20matters%20in%20groups.pdf.

Mehta, J. 2013. “How Paradigms Create Politics: The Transformation of American Educational Policy, 1980-2001.” *American Educational Research Journal* 50 (2): 285–324. https://doi.org/10.3102/0002831212471417.

Mehta, Jal. 2013. *The Allure of Order: High Hopes, Dashed Expectations, and the Troubled Quest to Remake American Schooling*. Oxford: Oxford University Press.

Merton, Robert C. 1973. “Theory of Rational Option Pricing.” *The Bell Journal of Economics and Management Science* 4 (1): 141. https://doi.org/10.2307/3003143.

Meseguer, Covadonga. 2004. “What Role for Learning? The Diffusion of Privatisation in OECD and Latin American Countries.” *Journal of Public Policy* 24 (03): 299–325.

Meyer, Alan, Peter J. Frost, and Karl E. Weick. 1998. “The Organization Science Jazz Festival: Improvisation as a Metaphor for Organizing—Overture.” *Organization Science* 9 (5): 540–42. https://doi.org/10.1287/orsc.9.5.540.

Meyer, Heinz-Dieter, and Brian Rowan. 2006. “Institutional Analysis and the Study of Education,” 13.

Meyer, John W., and Brian Rowan. 1977. “Institutionalized Organizations: Formal Structure as Myth and Ceremony.” *American Journal of Sociology* 83 (2): 340. https://doi.org/10.1086/226550.

Meyer, Peter. 2007. “Can Catholic Schools Be Saved?” *Education Next* 7 (2).

Mintzberg, Henry, and James A Waters. 1985. “Of Strategies, Deliberate and Emergent.” *Strategic Management Journal* 6 (3): 257–272.

Miron, Gary, and Brooks Applegate. 2007. “Teacher Attrition in Charter Schools.” East Lansing, MI: Great Lakes Center for Education Research and Practice. http://greatlakescenter.org/docs/Research/Miron_Attrition.pdf.

Nathan, Joe. 1999. *Charter Schools: Creating Hope and Opportunity for American Education*. 1st pbk. ed. The Jossey-Bass Education Series. San Francisco: Jossey-Bass Publishers.

National Center for Education Statistics. 2012. "Distance Education Courses for Public Elementary and Secondary School Students: 2009-10,." U.S. Department of Education. https://nces.ed.gov/programs/digest/d14/tables/dt14_218.20.asp.

———. 2014. "Conditions of Education 2014." U.S. Department of Education. http://nces.ed.gov/programs/coe/index.asp.

Neal, Derek. 2011. "The Design of Performance Pay in Education." National Bureau of Economic Research. http://www.nber.org/papers/w16710.

Nichols, Sharon L. 2007. "High-Stakes Testing: Does It Increase Achievement?" *Journal of Applied School Psychology* 23 (2): 47–64. https://doi.org/10.1300/J370v23n02_04.

Nichols, Sharon L., Gene V. Glass, and David C. Berliner. 2006. "High-Stakes Testing and Student Achievement: Does Accountability Pressure Increase Student Learning?" *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas* 14. http://www.redalyc.org/html/2750/275020543001/.

Nicolai, A., and D. Seidl. 2010. "That's Relevant! Different Forms of Practical Relevance in Management Science." *Organization Studies* 31 (9–10): 1257–85. https://doi.org/10.1177/0170840610374401.

Noda, Tomo, and Joseph L. Bower. 1996. "Strategy Making as Iterated Processes of Resource Allocation." *Strategic Management Journal* 17 (S1): 159–192. https://doi.org/10.1002/smj.4250171011.

Nolan, Kathleen. 2011. *Police in the Hallways: Discipline in an Urban High School*. Minneapolis: University of Minnesota Press.

Olberg, Amanda, and Michael J. Podgursky. 2011. "Charting a New Course to Retirement: How Charter Schools Handle Teacher Pensions." Thomas B. Fordham Institute. http://eric.ed.gov/?id=ED521685.

O'Reilly, Charles A., and Jennifer A. Chatman. 1996. "Culture as Social Control: Corporations, Culture, and Commitment." *Research in Organizational Behavior* 18: 157–200.

Pache, A.-C., and F. Santos. 2013. "Inside the Hybrid Organization: Selective Coupling as a Response to Competing Institutional Logics." *Academy of Management Journal* 56 (4): 972–1001. https://doi.org/10.5465/amj.2011.0405.

Payne, Charles M. 2008. *So Much Reform, so Little Change: The Persistence of Failure in Urban Schools*. Cambridge, MA: Harvard Education Press.

Peck, Craig, and Ulrich C. Reitzug. 2014. “School Turnaround Fever: The Paradoxes of a Historical Practice Promoted as a New Reform.” *Urban Education* 49 (1): 8–38.

Pérez-Nordtvedt, Liliana, G. Tyge Payne, Jeremy C. Short, and Ben L. Kedia. 2008. “An Entrainment-Based Model of Temporal Organizational Fit, Misfit, and Performance.” *Organization Science* 19 (5): 785–801. https://doi.org/10.1287/orsc.1070.0330.

Perrow, Charles. 2011. *Normal Accidents: Living with High Risk Technologies*. Princeton University Press.

Perry, Brea L, and Edward W Morris. 2014. “Suspending Progress.” *American Sociological Review* 79 (6): 1067– 1087.

Pew Research Center. 2019. “Internet/Broadband Fact Sheet.” https://www.pewinternet.org/fact-sheet/internet-broadband/.

Podgursky, Michael. 2006. “Teams versus Bureaucracies: Personnel Policy, Wage-Setting, and Teacher Quality in Traditional Public, Charter, and Private Schools.” *Education Working Paper Archive*. http://eric.ed.gov/?id=ED509018.

Podgursky, Michael, and Dale Ballou. 2001. “Personnel Policy in Charter Schools.” http://eric.ed.gov/?id=ED456567.

Podgursky, Michael J., and Matthew G. Springer. 2007. “Teacher Performance Pay: A Review.” *Journal of Policy Analysis and Management* 26 (4): 909–50. https://doi.org/10.1002/pam.20292.

Preston, Courtney, Ellen Goldring, Mark Berends, and Marisa Cannata. 2012. “School Innovation in District Context: Comparing Traditional Public Schools and Charter Schools.” *Economics of Education Review* 31 (2): 318–30. https://doi.org/10.1016/j.econedurev.2011.07.016.

Ramanujam, Rangaraj, and Paul S. Goodman. 2003. “Latent Errors and Adverse Organizational Consequences: A Conceptualization.” *Journal of Organizational Behavior* 24 (7): 815–36. https://doi.org/10.1002/job.218.

Rao, Hayagreeva. 1998. “Caveat Emptor: The Construction of Nonprofit Consumer Watchdog Organizations.” *American Journal of Sociology* 103 (4): 912–61. https://doi.org/10.1086/231293.

Reback, Randall, Jonah Rockoff, and Heather L Schwartz. 2014. “Under Pressure: Job Security, Resource Allocation, and Productivity in Schools under No Child Left Behind.” *American Economic Journal: Economic Policy* 6 (3): 207–241.

Renzulli, Linda A., and Vincent J. Roscigno. 2005. “Charter School Policy, Implementation, and Diffusion Across the United States.” *Sociology of Education* 78 (4): 344–66. https://doi.org/10.1177/003804070507800404.

Rice, Jennifer King. 2010. “The Impact of Teacher Experience: Examining the Evidence and Policy Implications. Brief No. 11.” *National Center for Analysis of Longitudinal Data in Education Research*. https://eric.ed.gov/?id=ED511988.

Rice, Jennifer King, and Betty Malen. 2003. “The Human Costs of Education Reform: The Case of School Reconstitution.” *Educational Administration Quarterly* 39 (5): 635–66. https://doi.org/10.1177/0013161X03257298.

Rios, Victor M. 2011. *Punished: Policing the Lives of Black and Latino Boys*. New Perspectives in Crime, Deviance, and Law Series. New York: New York University Press.

Rollwagen, John, and Donn McLellan. 1988. “Chartered Schools = Choices for Educators + Quality for All Students.” Minneapolis, MN: Citizens League. https://citizensleague.org/wp-content/uploads/2017/07/PolicyReportEducationNov-1988.pdf.

Rousseau, Denise M. 1985. “Issues of Level in Organizational Research: Multi-Level and Cross-Level Perspectives.” *Research in Organizational Behavior* 7: 1–37.

———. 2011. “Reinforcing the Micro/Macro Bridge: Organizational Thinking and Pluralistic Vehicles.” Edited by Herman Aguinis, Brian K. Boyd, Charles A. Pierce, and Jeremy C. Short. *Journal of Management* 37 (2): 429–42. https://doi.org/10.1177/0149206310372414.

Sarason, Seymour Bernard. 1971. *The Culture of the School and the Problem of Change*. Boston: Allyn and Bacon.

Schein, Edgar H. 2010. *Organizational Culture and Leadership*. 4th Edition. San Francisco, CA: John Wiley & Sons.

Schneider, Benjamin, Mark G. Ehrhart, and William H. Macey. 2013. “Organizational Climate and Culture.” *Annual Review of Psychology* 64 (1): 361–88. https://doi.org/10.1146/annurev-psych-113011-143809.

Sebring, Penny Bender, Elaine Allensworth, Anthony S Bryk, John Q Easton, and Stuart Luppescu. 2006. “The Essential Supports for School Improvement.” Chicago, IL: Consortium on Chicago School Research.

Seok, Junhee, H. Shaw Warren, Alex G. Cuenca, Michael N. Mindrinos, Henry V. Baker, Weihong Xu, Daniel R. Richards, et al. 2013. “Genomic Responses in Mouse Models Poorly Mimic Human Inflammatory Diseases.” Edited by Amer Abouhamze, Ulysses G. J. Balis, David G. Camp, Asit K. De, Brian G. Harbrecht, Douglas L. Hayden, Amit Kaushal, et al. *Proceedings of the National Academy of Sciences* 110 (9): 3507–3512. https://doi.org/10.1073/pnas.1222878110.

Shanker, Albert. 1988. “National Press Club Speech.” Washington, D.C., March 31. https://reuther.wayne.edu/files/64.43.pdf.

Simmons, Beth A., Frank Dobbin, and Geoffrey Garrett. 2006. "Introduction: The International Diffusion of Liberalism." *International Organization* 60 (04). https://doi.org/10.1017/S0020818306060267.

Simon, Jonathan. 2007. *Governing through Crime: How the War on Crime Transformed American Democracy and Created a Culture of Fear*. Oxford: Oxford University Press.

Smircich, Linda. 1983. "Concepts of Culture and Organizational Analysis." *Administrative Science Quarterly* 28 (3): 339–358.

Smith, Marshall S., Jennifer O'Day, and Susan H. Fuhrman. 1992. "State Policy and Systemic School Reform." *Educational Technology* 32 (11): 31–36.

Snellman, Kaisa E. 2012. "Window-Dressers and Closet Conformists: Organizational Decoupling Revisited." *Academy of Management Proceedings* 2012 (1): 1–1. https://doi.org/10.5465/ambpp.2012.301.

Stinchcombe, Arthur L. 1965. "Social Structure and Organizations." In *Handbook of Organizations*, edited by James G. March, 142–193. Chicago: Rand McNally.

Stoddard, Christiana, and Peter Kuhn. 2008. "Incentives and Effort in the Public Sector: Have US Education Reforms Increased Teachers' Work Hours?" *Economics of Education Review* 27 (1): 1–13. https://doi.org/10.1016/j.econedurev.2006.07.012.

Stuit, David A. 2010. "Are Bad Schools Immortal? The Scarcity of Turnarounds and Shutdowns in Both Charter and District Sectors." Washington, D.C.: Thomas B. Fordham Institute.

Stuit, David A., and Thomas M. Smith. 2012. "Explaining the Gap in Charter and Traditional Public School Teacher Turnover Rates." *Economics of Education Review* 31 (2): 268–79. https://doi.org/10.1016/j.econedurev.2011.09.007.

Teddlie, Charles, and Sam Stringfield. 1993. *Schools Make a Difference: Lessons Learned from a 10-Year Study of School Effects.* New York: Teachers College Press.

Thapa, Amrit, Jonathan Cohen, Shawn Guffey, and Ann Higgins-D'Alessandro. 2013. "A Review of School Climate Research." *Review of Educational Research* 83 (3): 357–85. https://doi.org/10.3102/0034654313483907.

The Center for Education Reform. 2011. "The State of Charter Schools: What We Know and What We Do Not about Performance and Accountability." Washington, DC: The Center for Education Reform. https://edreform.com/2011/12/charter-school-closure-report/.

Thernstrom, Abigail, and Stephan Thernstrom. 2003. *No Excuses: Closing the Racial Gap in Learning*. New York: Simon and Schuster.

Thompson, Charles L, Kathleen M Brown, Latricia W Townsend, Gary T Henry, and C Kevin Fortner. 2011. "Turning around North Carolina's Lowest Achieving Schools (2006-

2010).” Chapel Hill: Consortium for Educational Research and Evaluation—North Carolina. https://publicpolicy.unc.edu/files/2015/07/Turning-Around-NCs-Lowest-Achieving-Schools-2006-2010.pdf.

Thornton, Patricia H, William Ocasio, and Michael Lounsbury. 2012. *The Institutional Logics Perspective: A New Approach to Culture, Structure, and Process*. Oxford University Press.

TNTP. 2015. “The Mirage: Confronting the Hard Truth about Our Quest for Teacher Development.” TNTP.

Tolbert, Pamela S., and Lynne G. Zucker. 1983. “Institutional Sources of Change in the Formal Structure of Organizations: The Diffusion of Civil Service Reform, 1880-1935.” *Administrative Science Quarterly* 28 (1): 22. https://doi.org/10.2307/2392383.

Torres, A. Chris. 2016. “Is This Work Sustainable? Teacher Turnover and Perceptions of Workload in Charter Management Organizations.” *Urban Education* 51 (8): 891–914.

Trujillo, Tina, Janelle Scott, and Marialena Rivera. 2017. “Follow the Yellow Brick Road: Teach For America and the Making of Educational Leaders.” *American Journal of Education* 123 (3): 353–91. https://doi.org/10.1086/691232.

Turner, Barry A. 1976. “The Organizational and Interorganizational Development of Disasters.” *Administrative Science Quarterly*, 378–397.

Tushman, Michael L., and Philip Anderson. 1986. “Technological Discontinuities and Organizational Environments.” *Administrative Science Quarterly* 31 (3): 439. https://doi.org/10.2307/2392832.

U.S. Department of Education Office of Inspector General. 2016. “National Assessment of Charter and Education Management Organizations.” Washington, D.C.: U.S. Department of Education. https://www2.ed.gov/about/offices/list/oig/auditreports/fy2016/a02m0012.pdf.

Vaughan, Diane. 1997. *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*. University of Chicago Press.

———. 1999. “The Dark Side of Organizations: Mistake, Misconduct, and Disaster.” *Annual Review of Sociology*, 271–305.

Wang, Catherine L., and Pervaiz K. Ahmed. 2007. “Dynamic Capabilities: A Review and Research Agenda.” *International Journal of Management Reviews* 9 (1): 31–51. https://doi.org/10.1111/j.1468-2370.2007.00201.x.

Weber, Klaus, Gerald F. Davis, and Michael Lounsbury. 2009. “Policy as Myth and Ceremony? The Global Spread of Stock Exchanges, 1980–2005.” *Academy of Management Journal* 52 (6): 1319–1347.

Weeks, John. 2004. *Unpopular Culture: The Ritual of Complaint in a British Bank*. Chicago: University of Chicago Press.

Weick, Karl E. 1976. "Educational Organizations as Loosely Coupled Systems." *Administrative Science Quarterly* 21 (1): 1. https://doi.org/10.2307/2391875.

Westphal, James D., Ranjay Gulati, and Stephen M. Shortell. 1997. "Customization or Conformity? An Institutional and Network Perspective on the Content and Consequences of TQM Adoption." *Administrative Science Quarterly* 42 (2): 366. https://doi.org/10.2307/2393924.

Wharton, Amy S. 2009. "The Sociology of Emotional Labor." *Annual Review of Sociology* 35 (1): 147–65. https://doi.org/10.1146/annurev-soc-070308-115944.

Willis, Paul E. 1977. *Learning to Labor: How Working Class Kids Get Working Class Jobs*. Columbia University Press.

Wilson, JQ, and GL Kelling. 1982. "Broken Windows." *Atlantic Monthly* 249 (3): 29–36.

Wilson, Steven F. 2009. "Success at Scale in Charter Schooling." *Education Outlook*, March 2009. https://www.aei.org/publication/success-at-scale-in-charter-schooling/.

Wong, Kenneth K., and Francis X. Shen. 2003. "Measuring the Effectiveness of City and State Takeover as a School Reform Strategy." *Peabody Journal of Education* 78 (4): 89–119. https://doi.org/10.1207/S15327930PJE7804_06.

Worp, H. Bart van der, David W. Howells, Emily S. Sena, Michelle J. Porritt, Sarah Rewell, Victoria O'Collins, and Malcolm R. Macleod. 2010. "Can Animal Models of Disease Reliably Inform Human Studies?" *PLOS Medicine* 7 (3): 1–8. https://doi.org/10.1371/journal.pmed.1000245.

Wujec, Tom. 2010. *Build a Tower, Build a Team*. TED. https://www.ted.com/talks/tom_wujec_build_a_tower.

Yoon, Kwang Suk, Teresa Duncan, Silvia Wen-Yu Lee, Beth Scarloss, and Kathy L. Shapley. 2007. "Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement. Issues & Answers. REL 2007-No. 033." *Regional Educational Laboratory Southwest (NJ1)*. https://eric.ed.gov/?id=ED498548.

Zimmer, Ron, and Richard Buddin. 2009. "Is Charter School Competition in California Improving the Performance of Traditional Public Schools?" *Public Administration Review* 69 (5): 831–45. https://doi.org/10.1111/j.1540-6210.2009.02033.x.

Zimmer, Ron, Brian Gill, Kevin Booker, Stephane Lavertu, Tim R. Sass, and John Witte. 2009. "Competitive Effects of Charter Schools on Student Achievement in Traditional Public Schools." In *Charter Schools in Eight States: Effects on Achievement, Attainment,*

Integration, and Competition, 77–82. RAND Corporation. http://www.jstor.org/stable/10.7249/mg869bmg-joy-wpf.12.

Zucker, Lynne G. 1987. "Normal Change or Risky Business: Institutional Effects on the 'Hazard' of Change in Hospital Organizations, 1959-79." *Journal of Management Studies* 24 (6): 671–700. https://doi.org/10.1111/j.1467-6486.1987.tb00468.x.