Organizing for Quality in Education:
Individualistic and Systemic Approaches to Teacher Quality

by

Seneca Rosenberg

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Educational Studies)
in the University of Michigan
2012

Doctoral Committee:

Professor David K. Cohen (Chair)
Anthony S. Bryk, President of the Carnegie Foundation for the Advancement of Teaching
Professor Magdalene Lampert
Professor Carla O'Connor
Professor Kathleen Sutcliffe

2012

ACKNOWLEDGEMENTS

I would never have decided to seek my Ph.D. in education had it not been for my experience as a

new teacher in San Jose, California. I left the classroom with the firm conviction that students,

and particularly poor and minority students, deserve access to reliably strong opportunities to

learn, and that we must better understand how that might be achieved beyond individual

classrooms or schools. I am indebted to the students, colleagues, and mentors who inspired me to

engage in this work.

I was able to pursue graduate studies due to wide-ranging support—financial, intellectual,

and personal—from many different sources. The University of Michigan's Rackham Graduate

School and School of Education both provided substantial funding for my graduate work and for

the research and writing of this dissertation. I am grateful to both.

This dissertation was possible because Achievement First's leadership graciously granted

me permission to study the network. Their decision to do so is an example of AF's commitments

to transparency, collaboration, and learning with the goals of helping their students to achieve

and inspiring broader reforms—values that I found were reflected across many of the interviews

and observations I conducted. I am also thankful for the many staff members throughout the

network who were generous with their time and ideas, and who welcomed me to their network

events, schools, and classrooms. Some of these individuals are named in the body of the

dissertation while others must remain anonymous, but I deeply appreciate the contributions of all

who participated. I am particularly grateful to the two thoughtful principals who helped me to

plan my visits to their schools and who, along with their academic deans, were wonderful hosts

TABLE OF CONTENTS

LIST OF APPENDICES

ABSTRACT

Organizing for Quality in Education:
Individual and Systemic Approaches to Teacher Quality

by

Seneca Rosenberg

Chair: David K. Cohen

Reliability in education and its outcomes has long been elusive in the United States, especially for disadvantaged students. Because teachers are the most important school-based influence on student learning, this dissertation investigates the key challenges involved in defining and organizing to improve teaching quality in the U.S.

In the first of three essays, I examine the research literature on teacher and teaching quality from the 1890s to the present to understand why such sustained empirical work has failed to generate a coherent, actionable knowledge base for teaching. I identify five lines of research on teacher and teaching quality with different implications for reform. I conclude that two fundamental features of the American educational landscape—the absence of a common set of goals and technologies for practice, and a political system that encourages variability—have undermined their success.

The second and third essays are based upon a 2009-2010 case study of Achievement First (AF), a high-performing charter management organization that has built an alternative to a traditional school district. I draw upon 41 semi-structured interviews with network leaders,

school leaders, and teachers; archival data; and observations to investigate how AF was working to define, develop, and coordinate teaching quality throughout their rapidly expanding network.

I found that AF was in the early stages of constructing an infrastructure of practice—or set of integrated tools, structures, and practices—that was being used to develop the knowledge and skill of individual educators and their collective capability. This emergent infrastructure served as a scaffold for individual and organizational learning, and as a framework organizing the outcomes of these efforts; as a safety net for individual and organizational performance; and as a haven from the broader system. Particularly for more veteran educators, it could also act as an impediment to growth.

Even in its nascent form AF's infrastructure offers powerful insight into the possibilities and challenges involved in building a professional knowledge base for teaching. Still, the AF case study calls into question education reform strategies that assume that high-performing CMOs like AF can expand at the meteoric pace that would be necessary to reduce achievement gaps nationally.

CHAPTER ONE

INTRODUCTION

The U.S. system of education is not designed to support reliability in educational quality. Indeed, it has frequently been described as uncoordinated, with authority fragmented across multiple levels and branches of government (Cohen & Spillane, 1992; Elmore, 2004; Meyer & Rowan, 1978). A weak central power has supported the development of a strong tradition of local control in the financing and governance of schools throughout the country and has contributed to vast disparities in local spending and student achievement outcomes. This decentralization has also permitted nongovernmental groups to assert their influence on educators, whose responses to these pressures are idiosyncratic. In this environment, teachers—the most important school-based influence on student learning—vary widely in their effectiveness in helping students to learn (Hanushek & Rivkin, 2006).

This variability is problematic if the aims of public education include providing students with consistent access to high quality opportunities to learn. Its consequences can be especially detrimental for students whose families do not have the resources to help them achieve despite inadequate schooling. In this context, what might it take to provide reliably high-quality instruction in U.S. schools?

This dissertation includes three essays that examine some of the key challenges involved in defining and improving teaching quality in the U.S. In the first essay (Chapter Two), I examine the research literature on teacher and teaching quality from the last century to

understand why such voluminous and sustained empirical work has largely failed to generate a coherent research base around which consensus for large-scale teaching reforms might be built. Looking across these studies, I identify five categories of research on teacher and teaching quality, each with different implications for reform. The five categories treat teacher quality as a function of (a) the characteristics or qualifications of the teacher; (b) teacher behaviors; (c) social interaction, knowledge, and cognitive processes; (d) the organizations or systems of which they are a part; or (e) simply gains in students' achievement. I conclude that two fundamental features of the American educational landscape—including the absence of a common set of goals and technologies for practice, and the related fact that the U.S. educational system is tied to a political system which frequently frustrates efforts to develop such shared tools—have undermined the success of each of these five lines of inquiry, and will continue to do so until these features are acknowledged and addressed.

The second and third essays in this dissertation are based upon a 2009-2010 case study of Achievement First (AF), a high-performing charter management organization that has taken advantage of this decentralized system and its schools' relative autonomy as charters to build an alternative to a traditional school district. This study is draws upon 41 semi-structured interviews with AF leaders, school leaders, and teachers; archival data; and observations to investigate how AF's leaders and educators were working to define, develop, and coordinate teaching quality throughout their rapidly expanding network.

More specifically, the second essay (Chapter Three) examines the design of this professional support and management system, or "infrastructure of practice" (Cohen, 2011), and the ways that the elements of this infrastructure were intended to be used to sustain greater reliability in teaching quality throughout AF. I found that the network was creating robust

structures and practices organized into coherent systems that simultaneously developed the knowledge and skill of individual teachers and teachers' collective capability. In addition, I argue that combining individualistic with systemic approaches offered opportunities for supporting reliability in teaching quality at a scale rarely found in traditional school systems in the U.S. However, I also describe several of the internal and external challenges facing AF in this work that may curb their successes over time, particularly as the organization grows.

The third essay (Chapter Four) draws upon the same data and methods to investigate how this infrastructure was used and experienced by practitioners within the network, and to examine in greater depth the opportunities and challenges of building such a system in the U.S. context. I found AF's developing infrastructure worked in four primary ways within the network. First, it functioned as a framework and scaffold for individual and collective learning and practice, and as an outcome of these efforts. Second, tools, systems, and practices that were part of the infrastructure frequently acted as a kind of safety net for individual and organizational performance within AF by fostering the processes of "mindfulness" that characterize high-reliability organizations (Weick & Sutcliffe, 2001). Third, AF's infrastructure offered many of AF's staff members a shelter or haven from some of the dysfunctions of the broader educational system. In each of these ways, the AF infrastructure generated a host of social and technical resources that helped to sustain a minimum standard of instructional quality and provided several mechanisms through which educators, schools, and the network might learn and grow beyond that. Despite these many strengths, I found that some educators in AF—and particularly more veteran teachers and school leaders—experienced aspects of the infrastructure as impediments to individual and organizational learning over time in addition to understanding them as supports.

Some within the network also questioned the repercussions that these limitations might have for the AF's realization of the highest aspirations they held for their students.

Taken together, the essays are intended to inform the debate on teacher quality reform while illustrating the impressive social, political, and technical possibilities and challenges that taking on such work directly may incur—particularly when aimed at creating educational systems that might provide traditionally disadvantaged student populations with access to reliably strong learning opportunities across classrooms, schools, and academic years. In addition, the AF case study suggests that school systems may have a great deal to learn from this emergent network about how they might reorganize to support and enable better and more reliable student outcomes across classrooms and schools, and to learn from their efforts. However, the AF case study also identifies some areas of possible tension for the organization and calls into question the wisdom of education reform strategies that assume that high-performing charter school networks like AF will be able to expand at the meteoric pace that would be required to enroll large proportions of currently underserved students and to substantially improve their social and academic outcomes.

References

Cohen, D.K. (2011). *Teaching: Practice and its Predicaments*. Cambridge, MA: Harvard
University Press.

Cohen, D.K. and Spillane, J. (1992). Policy and practice: The relations between governance and
instruction. In G. Grant (Ed.), *Review of Research in Education, 18* (pp. 3-49).
Washington, DC: American Educational Research Association.

Elmore, R. (2004). *School Reform from the Inside Out: Policy, Practice, and Performance*.
Cambridge, MA: Harvard Education Press.

Hanushek, E.A. & Rivkin, S.G. (2006). Teacher Quality. In E.A. Hanushek & F. Welch (Eds.),
Handbook of the Economics of Education, Volume 2 (pp. 1052-1078). Amsterdam:
Elsevier North-Holland.

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

Weick, K.E., & Sutcliffe, K.M. (2001). *Managing the Unexpected: Assuring High Performance
in an Age of Complexity.* University of Michigan Pressing Problem Series. San Francisco:
Jossey-Bass.

CHAPTER TWO

WHAT MAKES A GOOD TEACHER?
PURSUING THE HOLY GRAIL OF EDUCATION RESEARCH

If effectiveness in teaching is unrecognizable and unmeasurable, it is also unpredictable. If effectiveness in teaching is either indeterminable or unpredictable, it indicates the terrific hazards of teaching as a vocation and the comparable hazards to society in obtaining a given quality of teaching service.

— *Arthur L. Odenweller, 1936*

It may surprise you—it was certainly surprising to us—but the field of education doesn't know very much at all about effective teaching. [...] This ignorance has serious ramifications. We can't give teachers the right kind of support because there's no way to distinguish the right kind from the wrong kind. We can't evaluate teaching because we're not consistent in what we're looking for. We can't spread best practices because we can't capture them in the first place.

— *Bill and Melinda Gates, 2011*

Since the turn of the twentieth century, the question "What makes a good teacher?" has captivated educational researchers who have produced thousands of studies in response. This pursuit has been driven largely by the conviction that a clear set of answers would provide a scientific basis for solving a host of practical problems in education. These practical problems include many of those at the core of the educational enterprise, such as how to identify the content and pedagogy that should be used in classrooms; specify the content, methods, structure, and evaluation of teacher education and professional development programs; assess teachers' practice, offer supports for their improvement, and make decisions about their tenure or dismissal; or develop policies that more equitably distribute effective teachers. Concerns about the returns on the public's substantial investment in teacher compensation provide further motivation. In the 2008-2009 school year, public schools in the United States employed over 3.2

million teachers, whose salaries and benefits absorbed approximately 55% of the 519 billion dollars spent on public elementary and secondary education nationwide (Johnson, Zhou, & Nakamoto, 2011; U.S. Department of Education, 2011).

Nonetheless, after more than a century of research, there remains a great deal of uncertainty among members of the education community about the aspects of teacher quality that are most important, how to measure them, or how to translate findings into a coherent knowledge base that might inform widespread practice improvement (Cochran-Smith, 2005; Gates Foundation, 2010; Goe, 2007; Hanushek & Rivkin, 2006; Hess, Rotherham, & Walsh, 2004; Rice, 2003). Researchers agree that there exists wide variation in teachers' effectiveness as measured by their unique contributions to their students' achievement gains, and that these deviations are large and meaningful for students' academic trajectories (e.g. Chetty, Freidman, and Rockoff, 2011; Hanushek & Rivkin, 2006; Rivkin, Hanushek, & Kain, 2005; Sanders & Rivers, 1996). However, they have been less successful in conclusively identifying the characteristics or practices that reliably discriminate between teachers who are more or less effective. Though policymakers and administrators often sidestep these issues by using proxy measures of teacher quality, including certification status, attainment of a master's degree, and additional coursework in content or method, the empirical record on these and other "quality" indicators remains quite mixed limiting their utility for administrators and policymakers (see, for example, Allen, 2003; Clotfelter, Ladd, & Vigdor, 2007; Darling-Hammond, 1999; Glazerman, Mayer, & Decker, 2006; Kane, Rockoff, & Staiger, 2006; Phillips, 2010).

To explore why these research efforts have generated only a fragmented knowledge base for teacher and teaching quality, I surveyed the research literature in teacher, and eventually teaching, quality from the turn of the last century when empirical efforts to explore teachers'

"efficiency" began in earnest.[1] These efforts shared the goal of understanding or defining teacher quality in ways that might improve schools' efficiency or equity. However, researchers in this area have conceptualized of and investigated teacher quality in very different ways with different implications for how it might be enhanced. Within this mass of studies, I identified five distinct categories of research on teacher and teaching quality to help make sense of these varied approaches and their implications for reform. The five categories treat *teacher quality as a function of* (a) the characteristics or qualifications of the teacher; (b) teacher behaviors; (c) social interaction, knowledge, and cognitive processes; (d) the organizations or systems of which they are a part; or (d) a teacher's estimated contribution to student achievement. The majority of research on teacher and teaching quality fits in one of these five categories, though there are some studies that usefully span more than one—including the high-profile Gates Foundation's Measures of Effective Teaching project, launched in 2009 with a $45 million dollar investment from the foundation—that I discuss after presenting the five core traditions.

In the first part of this essay, I describe each of these five lines of research and highlight the major aims and dilemmas facing researchers working within or across them. Given the considerable scope of this project, I deliberately focused attention on empirical work on teacher or teaching quality, or reviews of this literature. With the exception of a brief introduction to the environment in which these empirical studies were launched, I do not attempt to locate this literature in historical context; I also hold the studies apart from the larger theoretical debates within and among the academic disciplines in which they are situated. Though limited due to these omissions, this analysis of the leading traditions of research on the teacher and teaching

---

[1] A comprehensive review of the literature was not possible. By 1929, Charters and Waples report "an unwieldy mass of information" relevant to the training of teachers "too large for assimilation in a lifetime" (p. 3), and in 1974 Duncan and Biddle estimated that more than 10,000 studies on teacher effectiveness had been published. Since then, the amount of writing on the topic has continued to grow quickly. Therefore, I sampled the literature across time, using the content and reference lists of major studies and contemporary literature reviews to direct my search.

quality provides a useful overview of these efforts over the past century, and in doing so helps to illustrate some of the themes and challenges they share despite their seemingly vast differences.

In the second part of the essay, I look across these five lines of work and attempt to bring these commonalities into relief by exploring why these efforts have not resulted in a stronger consensus about teacher or teaching quality. While a number of possible explanations exist, I argue that two key features of the American educational landscape create the most fundamental problems undermining the success of each of these lines of inquiry. First, the United States' educational system lacks what Cohen (2011) has called an "infrastructure of practice" (p. 54), composed of the "extensive technical affordances that enable work" (p. 56). In education, this infrastructure would include a common set of exams, aligned with common curricula or curriculum frameworks around which teacher education could be designed; a specialized vocabulary of practice that could be used to communicate with others about the work and its improvement; and standards or norms of practice. Second, the U.S. educational system remains embedded in a political system that has thwarted the development of this educational infrastructure and has multiplied the uncertainties of absent goals and technologies (Cohen & Spillane 1992; Elmore, 1996; Rowan, 2002). I argue that both of these aspects of the educational context have strongly influenced the ways researchers have understood and investigated teacher quality, the problems they face in this work, and the extent to which their findings are useful for provoking meaningful instructional improvements.

In summary, three key questions guided my survey of the literature on teacher quality since the turn of the last century:

- How have researchers approached the study teacher quality in the United States, and what have these approaches yielded?

- What do these approaches suggest about how these researchers have understood teacher quality and how it might be improved?

- How have the features of the educational landscape in which this research has taken place influenced the studies, their findings, and how they might be used?

Below, I first briefly examine the environment in which intensive efforts to empirically investigate teacher and teaching quality began. Next, I present each of the five categories of research on teacher and teaching quality while emphasizing the major purposes and challenges of those working within them. Finally, I look across these traditions and explore the ways the U.S. educational context has influenced the success of individual studies as well as our ability to make use of their collective findings.

Launching the empirical study of teacher quality

Though interest in teacher and teaching quality had long been present in the United States and abroad, shifting needs and paradigms motivated a series of pointed inquiries into the topic by educators and researchers at the turn of the twentieth century. Soaring student enrollments, particularly at the high school level, were accompanied by a demand for greater numbers of teachers, the establishment of more and different types teacher training programs, and the growth of an administrative class to establish and oversee an increasingly large and complex system of education (Lagemann, 2000). Many of these "administrative progressives," as Tyack (1974) called them, were strongly influenced the ideas of Frederick Winslow Taylor. Taylor's ideas were first circulated widely through his 1895 essay advocating the "piece-rate system" of work that provided his contemporaries with a "pervasive metaphor for the practical value of defining all kinds of 'efficiency,' industrial, social, or educational, according to precise input-output correlations" (Lagemann, 2000, p. 79). These administrative progressives imagined that applying

such "scientific" approaches to the problems of education would result in greater efficiency in schools that had already been condemned by some observers as failing and decidedly "unscientific" (Rice, 1893, p. 6). In addition, for educational researchers situated in universities and administrators working in school systems, embracing this particular view of science held the promise of gaining status and legitimacy in the eyes of their peers in academic and professional circles (Lagemann, 2000, p. 21).

Within this environment, many educators and researchers trained their attention on what they believed to be a question of critical importance: "What makes an efficient teacher?" They reasoned that achieving its reply would provide them with the scientific basis and professional authority to solve, or justify their solutions for, a multitude of practical problems central to their work. Such problems included recruitment into and placement within the profession (Barr, 1945a; Barr, 1949; Barr, Torgerson, Johnson, Lyon & Walvoord, 1935; Boyce, 1915; Knight, 1922; Odenweller, 1936); teachers' promotion and retention (Barr et al., 1935; Boyce, 1915); and in-service supervision and "guidance" of teachers (Barr, 1940b; Barr et al., 1935; Boyce, 1915; Torgerson, 1934). Identifying effective teachers also promised to point researchers to the content and methods that should be used in schools (Odenweller, 1936; Thorndike, 1929) and to the development of a corresponding curriculum and pedagogy in the rapidly expanding normal schools and teacher training programs (Charters & Waples, 1929; Meriam, 1906). More generally, these educators and researchers believed knowledge of the hallmarks of effective teachers would permit them to develop scientific policies and practices that would prevent "waste" in the administration of teaching and in the classroom (Cureton, 1937; Odenweller, 1935), and keep "politics and favoritism" from dominating the schools (Boyce, 1915, p. 10).

The U.S. educational system has changed greatly since these early studies on teacher efficiency were conducted. Enrollments have continued to rise and have become nearly universal among school aged-children. Waves of education reform have attempted to make schools more relevant to students' lives, more academic in terms of their content, or more equitable in their outcomes, each falling short of their goals but leaving some imprint on the nature and structure of our school system. In the area of teacher quality, formal certification requirements, including minimum levels of educational attainment, rose dramatically and became centralized at the state level during the 1900s (Angus, 2001). Yet observers of the American educational system of the early twenty-first century will likely sympathize with the attention their predecessors lavished on the issue of teacher quality and the problems they hoped a scientific, or "research-based," approach to its study might allow them to solve. It is thus at this juncture in the history education research when the "objective" study of teacher quality began in earnest that I begin my essay on research on teacher and teaching quality.

Before I begin, however, a brief a note on the various terms that I and the scholars whose work I investigate use to refer to able teachers: They include quality, efficient, capable, good, effective, excellent, accomplished, and qualified. They also include terms like expert and ambitious, as well as phrases that denote normative expectations, like "should" or "ought" or "need" when referring to what teachers require for their success. Sometimes scholars and researchers use these terms interchangeably within a single study. In other cases, researchers announce they are quite deliberate in their choice of terms, avoiding, for example, words like "good" or "quality" because they seem to suggest some sort of value judgment rather than just the technical execution of goals set out by others. I have attempted to highlight these cases. Nonetheless, for the purposes of this essay, I have focused on the similarities underlying each of

these different ways of describing the teachers that researchers have, over the years, hoped to

identify or develop to improve the efficiency and equity of the American system of education.

Approaches to the study of teacher and teaching quality

I: Teacher quality as a function of the characteristics or qualifications of the teacher

Researchers working in the first of the five traditions of teacher quality research investigate the

extent to which teachers' characteristics (e.g. gender, age, intelligence, or personality) or their

qualifications (e.g. completion of professional training, years of experience, prior academic

achievement, or measured content knowledge) distinguish between better and worse teachers.

This type of research on teacher quality is appealing for a number of reasons. First, educators and

policymakers wish for clear, measurable indicators of teacher quality that will allow them to

make rational decisions about the teaching force. If research could demonstrate that particular

characteristics or qualifications are reliably associated with notably more effective teachers, this

would legitimize the use of these proxies in policy and practice and would suggest a clear path

for managing and improving teacher quality. In this framework, *teaching* quality is understood to

follow directly from *teacher* quality, thus attention to the former is not necessary. A second and

related reason research of this kind holds appeal is that much of the data necessary for its pursuit

can be quantified, collected through surveys or standardized tests, and analyzed with statistical

techniques to provide the appearance of objectivity so desired by many educators and

policymakers. Third, student gain scores on standardized, subject-specific tests, even when seen

as valid or desirable indicators of teachers' efficiency, are expensive to collect and calculate and

were even more so in the past (Boardman, 1928). What is more, information on students' gains is

unavailable for teachers entering the profession or teachers of subjects or grades for which such

exams are not regularly offered (Andrejko, 2004; Jones, 1946; Orleans, Clarke, Ostreicher, & Standlee, 1952). A set of more proximate criteria to direct personnel decisions is then useful.

Early studies of teacher efficiency

The majority of the research on teacher efficiency in the first half of the twentieth century treated teacher quality as a function of teachers' traits (Domas & Tiedman, 1950). A careful examination of the possibilities and frustrations these early researchers encountered in their work illuminates important parts of the terrain of research on teacher and teaching quality. Therefore, I devote more time to sketching out the landscape of this early work than I do to later lines of research.

Initiating the "objective" study of teacher quality: which traits matter?

When early researchers set out to identify the characteristics or qualifications associated with more efficient teachers, they began their studies in an environment they saw to be devoid of empirical work on the topic. In 1906, Junius L. Meriam, an adjunct professor of education at the University of Missouri credited with being the first to attempt the objective measurement of teacher efficiency (Boardman, 1928; Knight, 1922), published a set of studies in which he investigated whether or not teachers' professional training was linked with their subsequent success.[2] Prior to his study, he explained, a "scientific" lens had not been applied to these questions. "We have, therefore, only traditional standing and personal opinion to guide us," he lamented (p. 12).

Just a few years later, Ruediger and Strayer (1910) wrote, "The topic of the qualities of merit in teachers is so large and has been so little investigated (albeit frequently discussed) that a

---

[2] Specifically, he examined whether or not graduation from normal schools, teacher training schools, or no pedagogical schools, or the completion particular types of coursework seemed to "enter into the making of a capable teacher" (p. 51).

14

brief study like the present can do little more than to make a beginning and to raise questions" (p. 272). Therefore, they admit they "somewhat arbitrarily" (p. 273) selected teacher traits about which to gather data and to link them with teachers' efficiency. The characteristics and qualifications they chose included many that were the topic of subsequent studies: grade taught; certification and degree held; years of experience; health; personal appearance; strength of personality; ratings of teaching skill and method; the teachers' ability to keep order and to carry out suggestions; the accord between teacher and pupil; the degree of progressive scholarship or studiousness; and the "social factor" outside of school. In other early studies, teachers' intelligence, prior academic performance, or interest in school subjects were also measured (Boyce, 1915; Knight, 1922). While some of these attributes, and others such as teachers' height and weight or handwriting quality that were the focus of similar studies (see, e.g., Boyce, 1915; Odenweller, 1936) may seem absurd as potential indicators of teacher quality, it is useful to remember that many of these proposed proxies were merely guesses in the dark, and that their authors saw them as such.

The search for a criterion of teacher quality.

The most critical problem faced by early researchers working in this area was not related to the selection of possible proxies for teacher quality. Nor was it the measurement of the teachers' social, emotional, or psychological characteristics, though these did present a number of methodological difficulties with respect to validity and reliability that vexed researchers (see, e.g. Barr et al, 1935; Barr, 1940b; Boyce, 1915). Instead, the biggest challenge confronting these early researchers involved deciding on some criterion of quality against which these indicators might be tested. The Committee on the Criteria of Teacher Effectiveness (CCTE) (1952),

15

commissioned by the American Educational Research Association in 1950, defined a criterion as "a standard against which a measurement is made in estimating the validity of the measurement" (p. 242).[3] But in the measurement of teacher quality, what is the standard against which other instruments of measurement or prediction may be validated? The researchers of the first half of the twentieth century were well aware of this problem and its importance to their work. Meriam (1906) pointed to it:

> It must be frankly admitted at the outset that a strictly scientific treatment of the problem in hand [determining what contributes to teacher efficiency] is handicapped by the very nature of the data used. We have a strictly quantitative measure for land in the "foot-front" or acre, for coal in the ton or car-load. These are absolute measures and are universal. Not so in the measurement of scholarship or teaching efficiency. (p. 57)

Several decades later, in his review of the research on "The Measurement and Prediction of Teaching Ability," Torgerson (1934) wrote: "It is of course, obvious that valid instruments of measurement [to predict efficient teaching] cannot be determined until a valid and reliable criterion of teaching ability has been established" (p. 266). In her preface to *The Measurement of Teaching Efficiency*, Walker (1935) agreed, noting that without a clear criterion of teaching success, "none of the matters referred to [such as professional training and evaluation] can be studied with respect to their central function, the promotion of better teaching" (p. x). This led her to declare, "The lack of an adequate, concrete, objective, universal criterion for teaching ability is thus the primary source of trouble for all who would measure teaching" (pp. x-xi).

Absent a universal criterion for teaching ability a variety of criteria were proposed and used in these studies. For example, some of the earliest quantitative work in this area used *surveys of opinions* of "experts," "professionals," or students as the measure of quality against

---

[3] Members of this committee included many prominent researchers in the area, including A.S. Barr, Burley V. Bechdolt, Warren W. Coxe (who wrote an appended minority report to this 1952 report, and subsequently resigned from the committee), N. L. Gage, Jacob S. Orleans, H. H. Remmers (Chairman) and David G. Ryans.

which various proxies were tested. However, these efforts were quickly judged to be highly subjective and susceptible to ideological belief (Boyce, 1915; Crabbs, 1925; Torgerson, 1934). In an attempt to remedy some of these problems, researchers also developed *rating scales or scorecards* with which to judge teachers' effectiveness. Scorecards assigned a grade to teachers based on some set of items, while rating scales determined a teacher's position relative to his or her colleagues. Teachers were scored or ranked by their principal or supervisor (Meriam, 1906; Rudiger & Strayer, 1910; Odenweller, 1936), colleagues, students, or some combination of these (Knight, 1922; Boardman, 1928). Nonetheless, many observers complained that although these instruments were developed by "experts" their use remained highly subjective and they did not assess essential areas of teachers' work; had low or no correlations with student achievement gains; and were powerfully influenced by the "halo effect", or the overall judgment the rater made of the teacher as a person or colleague (Boyce, 1915; Odenweller, 1936).

Another type of criterion of teacher quality proposed were the scores of w*ritten tests of professional knowledge or skill* developed by experts as inexpensive "objective" alternatives (Boyce, 1915; Boardman, 1928). However, disagreement over the content of the tests and low correlations between their results and measures of student achievement caused many to doubt their utility. Others suggested that *teacher activities or behaviors* should serve as the benchmarks by which efficiency could be assessed. Yet many observers pointed out that judging teachers on their activities or behaviors could only be more objective than rating scales or scorecards after particular aspects of teacher practice had been linked with student learning (e.g. Torgerson, 1934; Odenweller, 1935; Orleans et al., 1952).

It may seem from this presentation that some agreement developed around the idea of using *student achievement or change* as the consensus criterion of quality, and to a certain extent

this was true. Although some scholars argued for investigations of changes in morality and character (Knight, 1922) or students' social adjustment or strength in cooperative planning (Orleans et al., 1952), many prominent researchers in the field argued vehemently in favor of using pupil gains in academic subjects as the key criterion of teacher quality.[4] Barr et al. (1935) explained why student change made sense as a consensus measure of efficiency:

> In selecting such a criterion of teaching efficiency one should consider the fact that the school exists to produce desirable changes in pupils. It is the function of the teacher to assist in the educative process by offering the pupil such guidance, direction, and assistance as will further the purposes for which the school exists. (p. 86)

In addition to appearing to be more objective than the rating scales or scorecards, assessing teachers on their students' achievement seemed to have a defensible theoretical rationale with respect to the aims of schools and their teachers.

However, a considerable subsection of researchers and other observers were deeply concerned by of the use of student academic gains as a criterion of teacher effectiveness and attacked the "objective" or "scientific" façade of achievement gains by pointing to two main types of problems with their use. The first set of problems was associated with technical problems involving the measurement of student gains. To begin with, many worried that available tests were only adequate for measuring students' acquisition of information and basic skill in school subjects rather than their development of deep understanding of the subject matter or in their abilities to apply their knowledge and skills (Orleans et al., 1952).[5] Using student gains on such tests as the benchmark of teacher quality might simply result in identifying "good

---

[4] See, for example, the work of A.S. Barr's work and that of many of his students, including L.E. Rostker, J.F. Rolfe, and C.V. La Duke; see also, Orleans, et al. (1952).

[5] Thorndike (1912) expressed little doubt that measurement of a broad range of educational outcomes would eventually be within reach. "There is no limit, theoretically, to the kind of thing for which scales are practicable," he espoused, including "the sense of evidence in history, excellence of judgment in affairs, devotion to the common good, or any quality, no matter how complex, that one may take" (p. 299).

drill masters who have no other desirable qualifications" (Walker, 1935; p. x). In addition, other

early researchers were unconvinced that there existed among them the sufficient technical ability

to isolate the effects of an individual teacher on his or her students' learning.[6] Although some

researchers examined associations between teachers' traits and raw gain scores of their students

in specific subjects (Brookover, 1945; Hill, 1921), others attempted elaborate controls of

students' intelligence and prior achievement using Accomplishment Ratios or Achievement

Quotients (Crabbs, 1925; Cureton, 1937), or later, regression models.[7] Nonetheless, concerns

about the difficulty of isolating individual teachers' effects on their students' learning led some

who believed that student gain scores were, theoretically, the best criterion of efficiency to

eschew them in their own work (Knight, 1922; Boardman, 1928).

A second set of concerns about using student achievement gains as a criterion of teacher

quality was linked to the absence of a broad consensus about the goals or objectives of education

in the United States. For example, some of these observers noted that a teacher who was

effective in helping her students realize progress towards goals that others judged worthless

could hardly be considered good, despite her efficiency. Absent widespread consensus on the

objectives and aims of education, researchers were left to their own devices to deal with this

problem. Walker (1935) made this point quite clearly, explaining that educators do not agree

with respect to what makes a good teacher or what good teachers do:

> This [lack of an objective criterion of teacher quality] would not be an insuperable
> difficulty if agreement could be reached concerning the desired outcomes of
> teaching, because then teaching could be measured not directly through

---

[6] See, for example, Odenweller (1936), who wrote, "The childlike faith in a simple measure of a teacher by the changes made in her pupils is shattered by the statistical complications in the methods and the inconsistencies in the findings. Staunch proponents of the immediate measurement of teaching by pupil achievement are many. But none of the students of the problem has said that even a fairly acceptable index of general teaching efficiency has yet been obtained by the measurement of pupil attainment" (p. 25).

[7] It is worth noting the challenge of doing many of these analyses without modern technology. Crabbs (1925) wrote that the procedure she used to estimate the change in AR for a sample of teachers in several subjects in a single district required 73,800 calculations, in addition to scoring 21,600 test papers and tabulating the results (p. 38).

measurement of the teacher but indirectly through measurement of pupil change. While the diversity of prevailing philosophies of education thus appears to be a bar to the construction of any universal measure of teaching ability, it may be possible to construct a measure useful within a specific frame of reference. (x)

Similarly, Gage and Orleans (1952) pointed out that without consensus on the aims of schooling, they were forced to consider effective teaching as local and idiosyncratic, with criterion "to be defined in terms of the ends adopted by any particular school or school system" (p. 296). Given these circumstances, they recommended that individual researchers clearly articulate the educational goals they are measuring effectiveness against. Yet they noted that regardless of the choices made by educational researchers each teacher under study may be aiming at different goals, making it difficult and perhaps unfair to compare or evaluate them.

Disappointing findings

Regardless of the characteristics and qualifications of teachers under study or the criterion of teacher quality selected—and even as statistical and measurement technologies at researchers' disposal became more sophisticated—the early studies of teacher efficiency are remarkably consistent in their findings. Researchers found, repeatedly, weak or no correlations between teacher traits and each of the criteria, or among the criteria (Barr & Jones, 1958).

Introducing his study correlating the ratings of teachers made by their colleagues, pupils, and supervisors, and linking these various ratings with a selection of teacher traits, Knight (1922) wrote: "As yet […], no one knows the exact formula for success in teaching. The complexity of personality and character and the many-sidedness of teaching have continually baffled useful analysis. We know that several measurable traits are *not* essential to successful teaching, but we do no know what traits *must* be present in superior instructors" (p. 4). Three decades later, Orleans et al. (1952) made a similar observation after reviewing studies that had tried to correlate

teacher characteristics with ratings "or with average measures of limited areas of pupil growth." They wrote, "An inspection of these many attempts leaves one with the feeling that, despite all the efforts expended, very little real progress has been made" (p. 641).

Some reacted optimistically to the paucity of findings and expressed confidence that improved instruments and methods would eventually reveal these relationships. This is especially understandable given that the field of measurement of social phenomenon was in its infancy in the first part of the 20th century. With technological advances, more resources, bigger sample sizes, etc., it seemed possible that science could deliver the "truth" of teacher effectiveness. Meriam (1906) hoped, "Time and experience may develop a standard of measurement of various mental traits, as the foot and ton, in physical measurements" (p. 58). Thorndike (1912) assured his followers "The changes that take place in intellect and character are coming to be measured with the same general technique, and, we may hope, with the same passion for clearness and precision, which has served the physical sciences for the last two hundred years" (p. 289).

However, not all researchers were quite as upbeat. The members of the CCTE (1952) concluded, "It is impossible to control, statistically or experimentally, all of the variables related to our dependent variables other than the specific independent variable under scrutiny. The variables are far too many, too complex, and too unmanageable for this kind of ideal ever to be realized" (p. 261). Orleans et al. (1952) decided that the main function of this line of work had been to demonstrate the overall futility of such an approach "a negative [lesson], but still perhaps a valuable one" (p. 641). Indeed, these early studies and their inability to uncover simple and straightforward relationships between teacher characteristics and qualifications and their success—however it was defined—helped to push scholars' thinking about what "effectiveness"

in teaching actually was, how it might be measured, and where it might reside; it also turned greater attention back to theory about the nature of teaching itself.

Out of the ashes—new directions for research on teacher quality

Out of these early research efforts, in fact, came a number of insights that helped to lay the groundwork for the other lines of research on teacher and teaching quality reviewed in this essay. For example, in interpreting the low correlations between their predictors and some measure of teacher quality, many investigators pointed to the complexity of teaching as a reason that single indicators may not reliably signal effectiveness, however they defined it. Therefore, some researchers attempted to *combine* indicators, and while they had moderately more success it is unclear that this was more than an artifact of over-specified models (Barr, 1945b; Hellfritzsch, 1945; Ryans & Wandt, 1952).

More importantly, this "combination" method was also accompanied by musings about the social and contextual nature of teacher effectiveness. The notion that teaching is largely a social endeavor, and that establishing and maintaining relationships with students, parents, other teachers, administrators, and community members was part of a teacher's work had long been part of the discourse in the area of teacher effectiveness (Barr, 1945b; Brookover, 1945; Hughes, 1958; Torgerson, 1934), especially as the insights of sociologists and social psychologists began to trickle into the field. But the growing focus on the situational and social aspects of teaching led Barr (1952) to suggest that effectiveness in fact might *exist* in these relationships and in their particular context. "Much of the research on teacher effectiveness seems to proceed as if the qualities in question resided entirely in the teacher. This may or may not be true," he wrote. Instead, he proposed, "'Teacher effectiveness' may be essentially a *relationship* between

teachers, pupils, and the other persons concerned with the educational undertaking, all affected by limiting and facilitating aspects of the immediate situation" [emphasis in original] (p. 172). Others agreed and emphasized that as a result, teacher effectiveness must not be thought of as a single dimension or trait, but must instead be understood to be quite context specific (Orleans et al., 1952). For example, Barr et al. (1955) wrote: "The search continues for a single generalized pattern of qualities or behaviors that characterize good teachers, notwithstanding the possibility that differential studies of teachers teaching different subjects to different sorts of pupils, under different conditions, and for different purposes might prove worthwhile" (p. 266).

Questions of this sort led many to decry an absence of theory to guide these deliberations. Barr (1940a) bemoaned a lack of theory about the organization of human ability, while others critiqued its absence with respect to teaching efficiency and how it might work (CCTE, 1952, 1953; Orleans et al., 1952), or to the teaching process itself (Hughes, 1958; Bellack & Huebner, 1960). As we will see in the other four lines of research reviewed in this essay, some researchers changed their focus, methods, or designs in an attempt to address some of these concerns.

Teacher characteristics and qualifications, continued

The disappointing results of this early research, together with new developments in statistics, psychology, and the social sciences, raised fresh questions about how teacher quality might be best understood and studied, and, as we will see, researchers did begin to investigate the topic in novel ways. However, research attempting to link teachers' characteristics and qualifications to their effectiveness has continued to the present for a variety of reasons. First, as some of the early researchers had hoped, it seemed possible that as researchers' technical skills and resources advanced, some indicators of quality could be identified among them. Second, as the American

educational system matured, relationships that had not existed previously could arise. For example, although proxies like teacher certification might stay the same in name, what they signaled might change. Third, when aggregated to the school or district level, proxy measures of teacher quality serve a useful role for contemporary researchers seeking to document the vastly unequal distribution of teacher quality. For example, these researchers have established that poor and minority students are far more likely than their white or more affluent peers to be taught by teachers who have five or fewer years of experience, lack credentials, do not have a major or minor in the subject they teach, or who initially failed their licensure exams. On aggregate, these distributional patterns reflect the deeply rooted inequities of a system in which traditions of local governance and financing overlap with segregation by race and class, contributing to troubling disparities in student achievement (Almy & Theokas, 2010; Boyd, Lankford, et al., 2008; Clotfelter, Ladd, & Vigdor, 2006; DeAngelis, White, & Presley, 2010; Peske & Haycock, 2006).

However, perhaps the primary reason that the search for teacher traits as proxies for quality continue is that the practical problems policymakers and educators need to solve—how to best recruit, train, reward, and allocate teachers—remain, as does the convenience of using those traits that can be easily measured for making such decisions. Though there have been some advances in the validity and usefulness of measuring teachers' contribution, or value-added, to their students' academic achievement gains in recent decades, there are still many technical challenges to this approach (see, e.g. the section of this essay on teacher quality as a function of student achievement), and this information is unavailable for teachers entering the profession or teachers of subjects or grades for which standardized exams are not consistently offered. Therefore, it is unsurprising that proxies of teacher quality are commonly used by policymakers to set standards for entering the teaching force or by administrators to differentiate between

applicants for teaching positions and to move teachers up a district's pay scale. The No Child

Left Behind Act (NCLB) of 2002 is a prominent example of this use of proxies. The federal

legislation required that by the 2005-2006 school year, all teachers in core academic areas be

"highly-qualified"— defined as holding a bachelor's degree; a state teaching credential; and

demonstrated subject matter competency (usually involving a state-approved test or a major or

minor in the subject) (US Congress, 2001).

Yet as was the case in the earlier studies, the empirical record on the ability of these and

other "quality" indicators to predict success in teaching is largely disappointing, with somewhat

mixed results. Even when researchers agree to use student learning gains as the criterion of

teacher quality, they tend to investigate the relationship between quality indicators and student

growth for different subject areas, at different grade levels, with different populations of

students, using different assessments, at variable units of analysis, to examine teachers using a

variety curricula with a diverse set of aims (see, e.g. Allen, 2003; Clotfelter, Ladd, & Vigdor,

2007; Darling-Hammond, 1999; Glazerman, Mayer, & Decker, 2006; Glazerman, Mayer, &

Decker, 2006; Goldhaber & Brewer, 2000; Kane, Rockoff, & Staiger, 2006; Philips, 2010;

Rowan, Correnti, & Miller, 2002).

Despite these challenges, there have been several recent efforts to synthesize and

interpret research in this area that have yielded conclusions that often conflict, particularly with

respect to the value of teacher certification and pedagogical coursework (e.g. Allen, 2003;

Darling-Hammond & Youngs, 2002; Walsh, 2001a; Wayne & Youngs, 2003; Wilson & Floden,

2003).[8] Similar debates in the literature rage with respect to the value other teacher traits,

---

[8] Walsh (2001a) concluded, for example, that she found "no credible research that supports using the teacher certification process as a regulatory barrier to teaching," and asserted, "Much of the research that is cited in support of certification reflects a level of scholarship that would not be tolerated in other professions" (p. 5). Darling-Hammond and Youngs (2002) rejoin: "Walsh's report excludes much of the evidence on the topic, misrepresents

including the utility of obtaining a master's degree, taking specific content or methods coursework, or matching students to teachers of the same race. There is more consensus that experienced teachers are more effective than novices; that teachers' academic or verbal ability tends to be linked with their effectiveness; and that in high school mathematics and perhaps science, teachers' subject-specific degrees and certification are positively related to student achievement, but even some of these findings are subject to debate (Wayne & Youngs, 2003).

As researchers, educators, and policymakers struggle to make sense of these often inconclusive or mixed results, many of their observations about the nature of this research echo the frustrations and hopes of their predecessors from the first half of the twentieth century. Many worry about the limitations of the technology employed these analyses. For example, in addition to her concerns about isolating the effects of teachers and their traits from other possible influences on students' achievement, Goe (2007) suggests that some existing relationships might not be detected because of small sample sizes or problems with the sensitivity or precision of the measurement tools and statistical analyses employed. She is hopeful, as were many of her forerunners, "that even better data systems and more precise statistical methods will be developed in the future" (p. 43). Goe also calls for an examination of teacher effectiveness within more specific contexts, suggesting that it makes sense to ask questions like "Within a given context—say, an at-risk urban school—what are the qualifications and characteristics associated with teachers who are effective at producing student achievement?" (pp. 44-45). Attentive to the findings with respect to advanced degrees and certification that vary by grade

---

many research findings, makes inaccurate claims about studies that have examined the consequences of preparation, and uses a double standard in evaluating the research" (pp. 13-14). Their review of the literature leads them to conclude that empirical studies "that have examined the influence of teacher education and certification on student achievement have often found significant relationships between these measures of teacher expertise and student achievement" (p. 18). Walsh responded to many of these assertions in her 2001 "Teacher Certification Reconsidered: Stumbling for Quality. A Rejoinder."

level and subject area, Rice (2003) makes a similar point, noting that "the limited research regarding the different types of students and different subject areas precludes definitive conclusions about how specific teacher qualifications vary across the dimensions" (p. 51). She also highlights the importance of attending to context in interpreting the research: "This literature review demonstrates that it is inappropriate and potentially misleading to draw generalizations about the effect of teacher attributes across levels of education, subject areas, and types of students" (p. 51).

Rice (2003) notes that there are two other problems plaguing those who would attempt to make sense of the research literature in this area. First, she observed, "the meaning of many of the teacher attributes studied in this body of research varies greatly across time and place" (p. 51). For example, certification requirements that vary by state, and the content of teacher education programs frequently varies across and within institutions. Therefore, she writes, "This idea of 'varied treatments' undermines the ability to draw general conclusions about what exactly matters in terms of teacher preparation" (p. 52). Second, Rice notes her conclusions "are largely based on the reported statistical significance of estimates resulting from empirical research," and "an effect that is statistically significant could also be trivial in size, limiting its policy relevance" (p. 52). Studies and reviews that simply highlighted significant findings without attending to the relevance of their size might lead policymakers and educators in unproductive directions. Taken together, then, contemporary studies of teacher quality as a function of teacher traits still leave many of our most fundamental questions about teacher quality unanswered, and suggest that more studies of this kind will not alone offer their solution.

…

In sum, research that attempts to link measurable teacher characteristics and qualifications with some criteria of teacher quality is appealing because of its enormous potential for policy and practice in education. In this tradition, teacher and teaching quality are conflated, with the second being understood to follow directly from the first. From this perspective policy efforts to improve teacher quality would involve reforming recruitment and hiring practices to ensure that the "right" people made it into the teaching profession, and perhaps some targeted forms of teacher education that could develop particular traits in prospective teachers. However, the results from this line of research have frequently been inconclusive, inconsistent, or even flatly contradictory. Though the causes for this are likely multiple, some researchers began to wonder if this focus on qualifications and characteristics as a way to understand and improve teacher quality was misguided.

II: Teacher quality as a function of teacher behaviors

Researchers seeking to link teachers' qualifications and characteristics to their effectiveness reached largely inconclusive findings during the first half of the twentieth century. A great deal of reflection about these disappointing findings followed. In addition to questioning their technical ability to measure and model these relationships, researchers also turned their attention to questions of the nature of teacher and teaching quality itself. Surveying the literature on teacher effectiveness conducted in earlier decades, Orleans et al. (1952) suggested, "Perhaps a major weakness of educational research has been the failure to do the basic thinking which is needed to insure that the right questions are being asked and that sound planning is being done." They continued, "In dealing with a problem which is as abstruse and complex as that of teacher effectiveness much time must be devoted at first to reading and deliberation" (p. 648).

One outcome of this type of reflection, when paired with the insights offered from developing theory and methods in psychology, sociology, and anthropology, was a new focus on what transpired inside of the classroom. "Research in teacher effectiveness has frequently bypassed the fundamental problem of identifying or conceptualizing the teaching process," Bellack and Huebner (1960) observed. "The obvious link between criterion and predictor, the teaching act, has been ignored or relegated to secondary importance" (p. 257). Writing just a few years later, Medley and Mitzel (1963) added:

> Certainly there is no more obvious approach to research on teaching than direct observation of the behavior of teachers while they teach and pupils while they learn. Yet it is a rare study indeed that includes any formal observation at all. In a typical example of research on teaching, the research worker limits himself to the manipulation or study of antecedents and consequents of whatever happens in the classroom while the teaching itself is going on, but never once looks into the classroom to see how the teacher actually teaches or how the pupils actually learn. (p. 247)

This new focus on the teaching and learning processes occurring within classrooms became the basis for an enormous outpouring of research related to better understanding teacher—and now teaching—quality.

Promises of process-product research

One of the most important branches of scholarship on classroom processes is frequently called "process-product" research. At its most basic, research in this tradition seeks to identify process variables (most frequently, teacher behaviors or groups of behaviors) that influence product variables (almost exclusively measures of students' achievement gains measured with standardized tests). From this perspective, variations in teacher quality follow from variations in teacher behaviors or behavioral sets, and an effective teacher is one who demonstrates behaviors associated with higher than expected gains in their students' average achievement.

Although the few previous empirical studies which had attempted to link teachers' behaviors with some measure of their effectiveness had met with little or no success (Jayne, 1945; Medley & Mitzel, 1959), this focus on the teaching process inspired researchers of the 1960s and 1970s with renewed hope that scientific research would guide them out of the chaos and confusion of previous educational practice. This air of newfound optimism bore a remarkable similarity to the earlier wave of certainty that the study of teacher traits would provide the basis for a science of education. Again, researchers professed faith in the power of science to eliminate teaching's uncertainties. For example, in introducing their seminal book, *The Study of Teaching*, Duncan and Biddle (1974) proclaimed: "Our dream is of an educational system whose procedures are governed by research and by theories that are empirically based" (p. vii). They contended, "the activities of teaching are reasonable, natural, rational events. They have discoverable causes and effects" (p. 12). Understanding these causes and effects, or these process-product relationships, promised to unlock the mystery of teacher and teaching quality, and subsequently, many of the mysteries of education policy and practice that the earlier research on teacher traits had not. Again, much of the focus was on improving the teaching force through teachers' recruitment, training, and supervision (Gage, 1963; Medley & Mitzel, 1963).

Typical process-product studies were conducted in naturalistic classroom settings using standardized observation systems to catalogue a teacher's behaviors (and sometimes those of the students). Some researchers relied on high inference "rating systems" to capture information that required the observer to make judgments about the behaviors they observed; these systems most frequently aimed to assess the degree to which teachers demonstrated qualities such as "clarity of presentation," "enthusiasm," or "helpful to students" using a Likert scale of some kind (Rosenshine, 1970, p. 281). However, the majority of studies in the process-product tradition

used low inference or "category systems" that were designed to rely less on the perspectives of the observer. These systems were intended to offer a more objective picture of a teacher's behaviors by providing a count or tally of a specific type of behavior—such as offering praise to students, asking open or closed questions, reviewing lesson content to close the lesson, or involving multiple students in responding to a question—instead of a more subjective rating of things like the teachers' warmth (Rosenshine, 1970, p. 282). These counts were analyzed alone, in combination, or in sequence with other observed behaviors. Standard process-product work then aggregated these individual counts or sequences across the observation periods (i.e. days of instruction or class periods) for each teacher before examining the extent to which variation in the mean incidence of such behaviors was related to variation in student achievement (Shulman, 1986a, p. 10).

The majority of the process-product studies conducted during the 1960s through the early 1990s were correlational. This meant that despite attempts to control for the influence of a variety of "presage" variables (i.e. teacher characteristics and qualifications of the type studied in the early research on teacher effectiveness) and "context" variables (e.g. pupil formative experiences, knowledge, or beliefs, and information about the schools, communities, and classrooms), the potential for unmeasured and omitted variables to bias the relationships observed between teacher behaviors and student learning remained. However, a small number of researchers did attempt to craft experiments that would provide a basis for making causal claims about observed relationships. These experiments were designed to systematically vary specific teacher behaviors that were hypothesized to be related to students' learning based on the results of the earlier correlational studies, psychological theory, and teachers' beliefs about instruction in carefully controlled but real classroom settings. Beginning in the late 1970s, a variety of

experimental studies were underway in which at least one group of teachers was trained in new

instructional methods while another group was not (for reviews, see Brophy & Good, 1986;

Gage & Needels, 1989; Rosenshine, 1983; Rosenshine & Stevens, 1986).

By the early 1980s, there was some indication that this body of work, especially that

derived from the experimental studies, had begun to contribute to a growing consensus on the

behaviors that characterized quality teaching and teachers (see, e.g. Leinhardt, Zigmond, &

Cooley, 1981). For example, Rosenshine (1983) and Rosenshine and Stevens (1986) drew on the

findings of 13 experimental studies of teacher behaviors to generate a list of major functions of

successful, "systematic" teaching. They felt confident that "the results of these studies are

consistently positive and indicate that there are specific instructional procedures which teachers

can be trained to follow and which can lead to increased achievement and student engagement in

their classrooms" (Rosenshine & Stevens, 1986, p. 376). In fact, they felt those procedures

related to effective teaching had been successfully identified, and that researchers could turn

their attention to investigating the merits of different approaches to their dissemination.

In a high profile and quite comprehensive review of the literature on "Teacher Behavior

and Student Achievement," Brophy and Good (1986) are a bit more cautious about their

conclusions, noting that the experimental findings of the studies in this area were less consistent

and had weaker findings than the correlational studies (p. 360). Nonetheless, they concluded,

"the last 15 years have produced an orderly knowledge base linking teacher behavior to

achievement" and believed that "if applied with proper attention to its limits, this knowledge

base should help improve teacher education and teaching practice" (p. 365). They found that

two common themes were prevalent in the process-product findings. First, the studies suggested

"academic learning is influenced by the amount of the time that students spend engaged in

appropriate academic tasks" and, secondly, "students learn more efficiently when their teachers first structure new information for them and help them relate it to what they already know." This structuring is enhanced when teachers help students to "monitor their performance and provide corrective feedback during recitation, drill, practice, or application activities (p. 366).

Problems with process-product research

Other researchers were less convinced of the strength of this knowledge base, however. Darling-Hammond and Wise (1985) claimed that "although this line of research indicates that what teachers do in the classroom does effect students, claims that discrete sets of behaviors consistently lead to increased student performance […] have been undermined by inconsistent and often contradictory findings" (p. 327). Nuthall and Alton-Lee (1990) noted that different reviewers from the previous decade had interpreted this research literature in divergent ways. They also highlighted several of the contradictions of the knowledge base. For example, they observed that teachers' use of "closed questions," which feature in Rosenshine's (1983) review, had been identified as being positively related with higher achievement in several studies, but so had teachers' employment of "open" or "higher-order" questions (p. 551). Similarly, teachers' use of "review," which also featured prominently in Rosenshine's synthesis, met with mixed results in correlational and experimental studies (p. 551).

In addition to criticisms about the inconsistency of results, research in this tradition faced a series of concerns about its underlying assumptions, methodology, and overall usefulness. Brophy (1988), who championed the potential of process-product work to empower teachers by providing a professional knowledge base to draw upon in their practice, worried about the misuse of the process-product findings given their limitations. He stated that for several reasons,

"neither teacher-effects data nor any other scientific data can directly prescribe guidelines for practice" (p. 7). First, he observed that while "[s]cientific findings can identify effective ways to attain given sets of prioritized educational objectives, […] they cannot make decisions for educators about what the objectives should be or how they should be prioritized." Instead, he suggested, "These are policy decisions to be made on the basis of moral, social, and political values (p. 7).[9] Brophy argued that absent policy decisions sanctioning the low-level knowledge and skills measured efficiently by the standardized achievement tests used by the majority of process-product studies, the teacher behaviors that were identified as contributing to above average gains on these tests could not be presented as guidelines for achieving the range of objectives we might care about. Second, Brophy argued that process-product studies had produced more information about "the quantity of instruction (how much active teaching occurs) than about its quality (what forms it takes and how well it is implemented)" (p. 9). Finally, he noted that process-product studies tended to focus on the behavioral differences between the least successful teachers and their colleagues, rather than differentiating between outstanding and adequate teachers. Therefore, already average teachers who would like to improve their practice may not find much to support their efforts in the process-product literature.

Given these limitations on the application of the process-product findings, Brophy (1988) emphasized that their valid use "requires interpretation by educators who are knowledgeable about classroom functioning and mindful of the limitations and qualifications that must be placed on any guidelines induced from such research" (p. 16). Many observers agreed, arguing that eliminating important features of the environment for the sake of generalizability—including the

---

[9] This is why Brophy (1988) prefers the term "teacher-effects" to "teacher-effectiveness": "'Teacher effectiveness' is a broad term that has meaning only in reference to a set of prioritized educational objectives, and most educators would want to consider several other objectives besides achievement-test gain in defining and assessing teacher effectiveness (developing student interest in subject matter, fostering the personal adjustment and mental health of individual students, developing a prosocial, cooperative group atmosphere in the class, etc.)" (p. 7).

content and curriculum being taught, grade level, students' backgrounds and motivations, teachers' intentions, and the semantic and social contexts in which teacher behaviors occurred— meant that teachers were lacking precisely the kind of information that they needed to apply the results of the studies (see, e.g. Brophy & Good, 1986, p. 366; Clark & Lampert, 1986; Hiebert, Gallimore, & Stigler, 2002; Nuthall, 2004).

The methodology used in the majority of process-product work was also subject to a great deal of disapproval. These criticisms also frequently condemned the theory or assumptions underlying the methodological choices. Berliner (1979), for example, questioned the assumption of stability in teachers' practice that are often implicit in researchers' search for relationships between processes and products that are temporally distant, or separated by subject matter: "How could the number or percentage of teacher verbal communications coded as praise statements in November influence results on achievement test items given in May? […] How could anyone expect to discover a relationship between a variable such as time spent lecturing on ecology and achievement test items that measure dictionary usage?" (p. 122).[10] Nuthall & Alton- Lee (1990) charged that process-product researchers' desire to produce valid, reliable, and generalizable results required the use of large sample sizes and data reduction techniques that eliminated concerns about how individual students experienced the classroom (p. 553). Shulman (1986a) described most process-product work as "unabashedly empirical and nontheoretical in tenor" because, though process-product research was able to establish that relationships between

---

[10] Some researchers in the process-product tradition shared this skepticism. Eyeing the debates about criteria of effectiveness that had plagued researchers in the early half of the century, Gage (1963) reasoned that "[r]ather than seek criteria for the over-all effectiveness of teachers in the many, varied facets of their roles, we may have better success with criteria of effectiveness in small, specifically defined aspects of the role" (p. 120), and coined the term "micro-criteria" to refer to them. Later, writing from the Stanford Center for Research and Development in Teaching, Gage (1968) demonstrated how well this concept complimented a "technical skills" and "micro-teaching" approach to teacher education, in which trainees engage in practice with teaching exercise scaled down in terms of time, class size, and task so the exercises can be recorded, evaluated, and re-taught in an improved manner (p. 121); this approach was also used to determine the relative importance and effectiveness of various teaching behaviors.

behaviors and outcomes existed, it was unable to determine "*why* particular combinations of teacher behavior led to gain and others did not, a question of theory" (Shulman, 1986a, p. 13).

The criterion of effectiveness almost always used in process-product studies—student achievement gains—was the focus of a good deal of the criticism aimed at this tradition of research as well. As we have seen, Brophy (1988) observed that much process-product research was only successful at identifying as effective those practices that were aimed at increasing students' low level knowledge and skills that could be measured with the existing standardized achievement assessments.[11] In addition, Shavelson, Webb, and Burstein (1986) offered a cogent analysis of four problematic assumptions underlying the use of the standardized, norm-referenced achievement tests so frequently used as the "product" in this research tradition.

First, they noted that effectiveness, as operationalized by process-product research, "assumes commonality of curriculum goals, objectives, and content coverage across classrooms because one standardized achievement test is used to judge the effectiveness in all classrooms" (p. 52). This assumption is problematic because standardized achievement tests and textbooks vary widely in the content they cover, and because the degree of correspondence between what is taught and what is tested is strongly related to students' test performance (p. 53). Therefore, without accounting for curricular variation, or students' opportunity to learn, "process-product research provides an inadequate basis for judgments about the plausible sources of performance gains that may be associated with teaching practices" (p. 54). Second, Shavelson et al. (1986) note that in most process-product work, "Effectiveness is strictly summative in its measurement of subject matter knowledge. It is not what students know or don't that matters, but the

---

[11] Duncan and Biddle (1974) worried about this possibility as well: "This [student achievement] is not only an insensitive variable, but it may be off mark for our purposes. Consider the finding that teacher use of higher cognitive demand leads to lower pupil achievement. It seems possible to us that lower cognitive demand is more efficient for putting across facts, while higher cognitive demand encourages independence of thought. The latter, of course, is not measured by standardized achievement tests" (p. 409).

accumulated quantity of their knowledge in comparison with students in other classrooms" (p.

52). These summary scores neither inform us about the *types* of questions students answer

correctly or incorrectly, nor about the changes in students' cognitive processes associated with

instruction (i.e. a student might answer the question correctly on both the pretest and posttest, but

draw on different cognitive structures or strategies while doing so).

Third, Shavelson et al. (1986) pointed out that in much process-product work,

"Performance on the effectiveness measure is equated with knowledge or skill in subject matter.

There is no notion of "less than best effort," guessing, partial knowledge, or test taking skill" (p.

52). The authors noted that if these "extra-knowledge" influences on students' performance are

ignored in the study design and analysis, then process-product research cannot claim to measure

the direct impact of teachers' practice on their students' knowledge (p. 56). Finally, Shavelson et

al. observe that in most process-product research, "Effectiveness is strictly aggregative across

students within classroom" (p. 52). Using the classroom as the unit of analysis and focusing on

students' mean gains can be problematic because this approach obscures data about within-class

variability in students' performance and experience and "thus misses evidence of a teacher's

differential effectiveness" (p. 56).

These critiques of process-product research, together with a precipitous decline in

funding for the expensive, large-scale observational studies and experiments that this tradition

requires contributed to a decline of the popularity of this approach in the late 1980s and the early

1990s.[12] However, research that attempts to identify a link between specific teacher behaviors or

practices with student learning continues (see, for example, Bodovski & Farkas, 2007; Milesi &

---

[12] From approximately 1965 through 1980, funding for observational studies in the United States was provide by a
variety of sources, including the federal government (Gage, 1994; Gage & Needels, 1989). The Office of Education
(OE) and the National Institute of Education (NIE) in particular supported large-scale process-product research
(Brophy & Good, 1986), including The Instructional Dimensions Study (see Cooley & Leinhardt, 1980) and
evaluations of Project Follow Through (see, e.g., Stallings, 1975).

Gamoran, 2006; Seidel & Shavelson, 2007; Taylor, Pearson, Peterson, & Rodriguez, 2003), and some would regard any research of this kind to be part of the process-product tradition. For example, Gage and Needels (1989) argued that there is nothing to prevent process-product researchers from using more sophisticated outcomes or process variables, models, or even from employing ethnographic methods in search of better insight into the mechanisms that connect process-product relationships (p. 291). According to their formulation, the essence of ongoing process-product work is the search for "knowledge of the connections between what teachers do and what students learn" (p. 295). They contend, "knowledge of such relationships is central to the understanding and improvement of teaching" (p. 294)—and that the search for these relationships remains, and should remain, strongly represented in current educational research.

<center>…</center>

In brief, often called "process-product" studies, research in this tradition focuses more on *teaching* than the *teacher*; in this case it was assumed that teacher quality followed from a teachers' use of particular behaviors or sets of behaviors that had could be linked to some criterion of teacher quality (almost always student achievement gains). This line of research assumes that policy aimed at improving teacher quality should focus on training teachers to enact behaviors with established relationships to desirable student outcomes. However, isolating causal relationships between specific teaching practices and student outcomes proved to be quite challenging, and the search for generalizable results often rendered the findings of such studies difficult to implement by excising the contextual features that might make them particularly useful for practitioners as they attempted to apply them to practice. Much of this work took place in the 1960s through the early 1990s.

III: Teacher quality as a function of social interaction, knowledge, or cognitive processes

Weak and inconsistent results from research that attempted to link teacher traits to their effectiveness during the early decades of the twentieth century helped to modify the focus of research on teacher quality. Together with advances in the theory and method of psychology and the social sciences, these small or inconclusive findings helped to direct researchers' attention to the teaching and learning processes that occurred *inside* the classroom. The process-product research tradition constitutes one approach to the study of this topic. This line of research has been successful in establishing that what teachers do with their students can influence their learning, and in identifying some promising general instructional practices. Nonetheless, many of the findings—especially those from the experimental process-product studies—were inconclusive and even contradictory, with frequently small relationships between observed teacher behaviors and student achievement gains.

Perhaps more importantly, many disagreed with the assumptions underlying process-product research. Some scholars rejected the behaviorist undercurrents of the process-product work that appeared to be "based on the beliefs that teaching consists of a repertoire of behaviors or teaching methods, and that student learning follows more or less directly from the frequency with which teachers use specific behaviors or apply a specific method" (Nuthall, 2004, p. 286). In addition, many of the process-product studies that involved behavioral "counts" seemed to assume that more of any behavior was always better, rather than attempting to determine optimal levels of the behavior, or how these optimal levels might co-vary by content area or grade level (Nuthall, 2004). Many of these studies also failed to develop theory to make sense of why behaviors might be connected to student outcomes (Shulman, 1986a). Other scholars fretted about the conception of the teacher that process-product research seemed to imply. "This

[process-product] approach emphasizes the actions of teachers rather than their professional judgments and attempts to capture the activity of teaching by identifying sets of discrete behaviors reproducible from one teacher and one classroom to the next," wrote Cochran-Smith and Lytle (1990, p. 2). The use of students' gains on achievement tests that were often unaligned with the classroom curricula, or that focused on basic skills rather than students' ability to analyze or apply what they had learned as the criterion of teachers' effectiveness troubled observers as well.

What is more, some scholars suggested that process-product researchers might be blind to a variety of other more fruitful lines of work that would better inform their understanding of the relationship between teaching and learning. For example, Nuthall (2004) wrote, "observations that resulted only in simple frequencies of behaviors did not record either the real significance of the behaviors or the meanings they might have for those who enacted or experienced them" (p. 285). Nuthall and Alton-Lee (1990) suggested "excessive concern with the objectivity and with the reliability of observation and measurement had led to the exclusion of the enormously rich sources of personal and individual experience available to researchers in the classroom" (p. 553).

Thus, a second contingent of researchers who focused their attention on what occurred inside the classroom did so in ways that attempted to address some of these critiques of process-product work and to make use of alternative theoretical and methodological approaches to the study of the teaching and learning process. While it is very difficult and possibly imprudent to wrangle this vast variety of studies into one category, I suggest that researchers in this tradition share an emphasis on the social and cognitive aspects of teaching and learning. Far from assuming that the same instructional stimulus will cause the same learning in students within and between classrooms, this line of work focuses on how these processes are mediated or

understood by individual students and teachers given their background, knowledge, experience, or abilities, often in highly variable, context and content specific ways. From this perspective, teacher quality involves a teacher's ability to identify, construct, and manage these aspects of the teaching-learning process, and to position students to learn. The improvement of teacher quality would thus require acting on the motivations, knowledge base, and social resources to which teachers have access in designing and conducting instruction. This line of work also has implications for ways in which educational systems might be organized to influence the resources a teacher is able to draw upon in his or her work—a topic that is taken up in the following section.

Social interaction, knowledge, and cognitive processes in the classroom

An important variation within process-product research questioned the direct link between teacher behaviors and student learning and the notion that teachers uniquely cause student learning. While maintaining the basic methods and structure of research in the process-product tradition, some researchers directed attention to *student* thoughts and behaviors in the classroom as potential mediators in the relationship between what teachers did and what students learned. Students' motivation, expectations, aims, time on task, task persistence, and interactions with peers were all tested as possible mediators of this relationship (Berliner, 1979; Cooley & Leinhardt, 1980; see also Doyle, 1977, Gage & Needels, 1989, Shulman, 1986a). Researchers accumulated evidence that students were not passive recipients of instruction but rather acted and processed information in diverse ways with variable consequences for their learning.

Others who probed this process more deeply looked beyond observable behaviors and attitudes measurable by survey instruments to the social and cognitive mediating mechanisms

involved in students' learning in the classroom. Shulman (1986a) identified two broad types of research of this kind. The first is grounded in a sociological tradition, and focuses on the social processes that influence the work teachers and students do together in the classroom, including the ways in which motivation and goals are developed and defined (p. 16). The second type is grounded instead in cognitive psychology, with a specific focus on students' apprehension of school subjects (p. 16), and demands a different view of students and teachers. Shulman suggested that by combining these two perspectives—the social from sociology and the cognitive from cognitive psychology—one can imagine the learner as making sense of both the classroom's social reality and the content of instruction simultaneously, given his or her prior experiences, available learning strategies or meta-cognitive processes, abilities or aptitudes, group memberships (including race/ethnicity and socioeconomic status), or the characteristics or qualities of their immediate setting (p. 17). This model suggests that improving teaching requires first understanding how teaching influences students' thinking, their apprehension of processes or strategies for learning and self-regulation (Seidel & Shavelson, 2007), and ultimately students' academic engagement and performance (Weinstein & Mayer, 1986; Wittrock, 1986).

Because students are not the only people doing cognitive work in the classroom, another group of researchers, many of whom had been trained in cognitive psychology, trained their attention not just on teachers' behaviors, but on their thinking and decision-making processes in addition to their behaviors (e.g. Feiman-Nemser & Floden, 1986). One group of these studies examined teachers' thinking as it relates to the kind of teacher behaviors observed in process-product studies. For example, in their prominent 1986 review of research on what was then a relatively recent focus on teacher thinking, Clark and Peterson reasoned that it is important to focus on teachers' thinking because these processes determine, in large part, their behaviors and,

42

subsequently, students' achievement (p. 255). They delineated three main categories of research on teachers' thought processes. First, research on teacher planning focused on the thought processes that occur before classroom interaction and after, in the form of teacher reflection. Second, inquiries into teachers' interactive thoughts and decisions focused instead on teacher thinking that takes place in the classroom during interactions with students; such thinking may cause teachers to revise their plans or behaviors while in the process of teaching. Finally, studies of teachers' theories and beliefs focused on "the rich store of knowledge that teachers have that affects their planning and their interactive thoughts and decisions" (p. 258). Together, these studies of teacher thinking suggested that teachers indeed play a crucial role in determining which aspects of the curriculum are translated into instruction as well as how time is allocated within the classroom, and that teachers' theories and beliefs can influence the how teachers interpret and enact the curricula and the events of their classrooms. These findings have potentially serious implications for those who would attempt to intervene to "improve" teaching.

Looking across the research on teacher cognition and process-product studies of teaching, Shulman (1986a, 1986b) and his colleagues drew attention to what they viewed as a distressing gap in the research. While some researchers had used the academic subject as a "context" variable, he worried they had largely neglected the study of teachers' understanding of subject matter and how this understanding related to instruction. Shulman (1986b) argued that neither content knowledge nor pedagogical knowledge divorced from content is adequate for teachers. Teachers also need what he termed "pedagogical content knowledge," or the "particular form of content knowledge that embodies the aspects of content most germane to its teachability" (p. 9). This includes a teacher's knowledge of useful ways to represent key topics of the content to students as well as an understanding of the typical conceptions or misconceptions that students of

various age levels and backgrounds might bring to the learning situation (p. 9). Wilson, Shulman, and Richert (1987) argued that pedagogical content knowledge is "enriched and enhanced by other kinds of knowledge—knowledge of the learner, knowledge of the curriculum, knowledge of the context, knowledge of pedagogy" (p. 114). It is also accompanied by pedagogical reasoning that supports teachers in constructing these representations of subject matter given the particulars of the classroom situation and attends to the normative aspects of teaching as well as to teaching methods (Shulman, 1987; Wilson et al., 1987).[13]

Many thoughtful studies have subsequently attempted to examine the ways in which subject matter might influence teaching (see, e.g., Stodolsky, 1998) and how teachers' understanding of subject matter relates to their instruction (see, e.g., Ball & Bass, 2003; Grossman, 1989; Lampert, 1986, 1990, 2001; Wineburg, 2001). In addition, a number of national organizations, including the National Reading Panel, the International Reading Association, the National Council of Teachers of Mathematics, and the National Board for Professional Teaching Standards have begun to take leadership roles in developing consensus documents in which current research knowledge is synthesized, or standards for teachers and teaching established, that emphasize subject matter knowledge for teaching.

Still other researchers have focused on their attention on additional aspects of what teachers need to know, and know how to do. For example, some studies have honed in on the "knowledge of the learner" aspect of teacher knowledge and have focused explicitly on excellence in the teaching of poor or minority students traditionally underserved by U.S. schools, sometimes as it is related to instruction in particular subject matter areas (e.g. Ladson-Billings,

---

[13] Discussions of teacher quality or of "good" teachers or teaching are rarely explicit in the writing of Shulman and his colleagues. However, they frequently use words like "expert" or "expertise", phrases that allude to what teachers "should know and know how to do", or ideas about "able teachers," "excellent teaching," or "ambitious teaching" that suggest some judgment with respect to the quality of the teachers or teaching.

1995a, 1995b; Gutstein, Lipman, Hernández, & de los Reyes, 1997; Lee, 2001). Ladson-Billings introduced the term "culturally relevant pedagogy" (CRP) to refer to one type of teaching that is attentive to poor and minority students' needs. According to Ladson-Billings (1995a), CRP includes three criteria: "(a) Students must experience academic success; (b) students must develop and/or maintain cultural competence; and (c) students must develop a critical consciousness through which they challenge the status quo of the current social order" (p. 160). Though her observation of teachers whose instruction attended to each of these criteria, Ladson-Billings (1995b) found that rather than exhibiting similarities in teaching behaviors or strategies, these teachers shared similar sets of beliefs about themselves and others; the importance of creating a system of social relationships that support them in working towards the three criteria of CRP; and dynamic and critical conceptions of knowledge. From this view, "good" teachers need to understand the role of culture and how it may function in education in addition to the other types of teacher knowledge articulated by her colleagues (p. 483).

Echoing some of the critiques of the kind of knowledge-base that might be generated from simplistic process-product research, Ball and Cohen (1999) observed to that in addition to these various kinds of teacher knowledge, teachers must know how to pursue more knowledge. In other words, they need to develop the capability to learn "in and from practice," or "how to learn in the contexts of their work" (p. 10). Although other types of knowledge of subject matter, children, and pedagogy can be important resources for practice, "practice cannot be wholly equipped by some well-considered body of knowledge" since "teaching occurs in particulars—particular students interacting with particular teachers over particular ideas in particular circumstances" (p. 10).

Regardless of the specific targets of their attention, researchers in this tradition, who view teacher quality as a function of social interaction, knowledge, or cognitive processes, required a different set of methodological tools to investigate their questions.

Methodological shift

The turn towards the social and cognitive aspects of the teaching-learning process was accompanied by a new set of methodological approaches and assumptions. Some topics of study were pursued with observational research methods similar to those used in the process-product tradition, with similar strengths and weaknesses. For example, some studies of students' actions and thoughts used observational tools, measurement instruments, and models quite similar to those used in process-product research. Similarly, some scholars of teacher thought quantified and described their data in analogous ways, and positioned their work as informing that of process-product research as well. However, they used data collection techniques quite different from those of process-product research in an attempt to uncover normally hidden sources of information on teachers' thoughts. These approaches included asking teachers to think aloud, "or verbalize all of his or her thoughts while engaged in a task" like lesson planning or evaluating teaching materials; engaging teachers in "stimulated recall" technique in which teachers are asked to listen to or watch a previously recorded teaching episode and asking the teacher "to recollect and report on his or her thoughts and decisions during the teaching episode"; or asking teachers to keep a journal in which they record their instructional plans and reasoning about them (Clark & Peterson, 1986, p. 259). Other forms of interviewing or intensive case studies offering description of students and teachers as they interacted in the classroom, or teaching portfolios including artifacts of practice have also frequently been used to gather data about classroom

interaction and student work (see, e.g., Evertson & Green, 1986; Junker, Weisberg, Matsumura, Crosson, Wolf, Levison, et al., 2006; Nuthall & Alton-Lee, 1990; Stodolsky, 1998; Wilson, et al., 1987).[14]

Other researchers relied on methodological approaches that Erickson (1986) joined together under the umbrella term *interpretive*. Based in part on research traditions from sociology, anthropology, and linguistics, these various methods involved somewhat different techniques for data collection and analysis, and rested on divergent assumptions. However, Erickson characterized interpretive studies in education as those that shared a common set of concerns: "(a) the nature of classrooms as socially and culturally organized environments for learning, (b) the nature of teaching as one, *but only one*, aspect of the reflexive learning environment, and (c) the nature (and content) of the meaning-perspectives of teacher and learner as intrinsic to the educational process" (p. 120). "Interpretive" scholars, therefore, understood the teaching-learning process in very different ways from their colleagues in the process-product tradition, or those who attempted to link simple characteristics or traits to teachers' effectiveness.

The minds and interactions of teachers and their students are important from the interpretive perspective because individuals decipher and make meaning of social life in the classroom, making choices and acting according to this reading of their situation. To understand

---

[14] One example of this comes from researchers at the Learning Research and Development Center at the University of Pittsburgh. Since 2002, these researchers have been working on the Instructional Quality Assessment (IQA), which aims to assess the quality of classroom instruction. The IQA is still in development, but currently involves direct observations of teacher behaviors and student tasks, the collection of samples of student work on teacher-generated assignments, and brief interviews with students and teachers (Junker, Weisberg, Matsumura, Crosson, Wolf, Levison, et al., 2006). The observations and assignments are evaluated using 4-point rubrics across several dimensions of quality, such as the academic rigor of lessons, the clarity of expectations, the extent to which self-management of learning is emphasized, and the use of accountable talk (in which members of the classroom community hold each other accountable for the accuracy and rigor of their ideas and communications). Because the authors hope to create a low-inference observation system with which to evaluate classrooms, they use quantitative indicators and checklists to steer raters' decisions on the rubrics. Eventually, the researchers plan on exploring the relationship between their instrument and its subcomponents with student achievement gains on a variety of types of standardized assessments (Junker, et al., 2006).

causal mechanisms in human social life, then, requires systematic, or objective, analysis of "subjective" meaning, which varies across individuals and groups, and across time. As a result, both formal and informal social systems of the classroom, in which individuals may have different statuses and roles, and of which participants may be more or less aware, are essential for an understanding the social ecology of the classroom. The success of students and teachers in this environment depends on their abilities to negotiate among the various demands of the classroom ecology (Doyle, 1977). From this view, Erickson (1986) explained, "teacher effectiveness is a matter of the nature of the social organization of classroom life […] whose construction is largely, but not exclusively, the responsibility of the teacher as instructional leader" (p. 133). This construction is rooted in the specific circumstances of the classroom and the individuals of which it is composed, but does not ignore the non-local influences of the cultural and social contexts in which they are embedded. Together with the insights about situative cognition, this outlook is carried over in many of the research studies reviewed in the next section of this essay, which examines teacher quality as a function of the organizations and systems of which they are a part.

Interpretive studies most frequently involve participant observational fieldwork, an approach that allows researchers to investigate the dynamic social organization of the environment in which these interpretations and meaning-making processes take place for specific individuals (Erickson, 1986, p. 128), and the ways in which non-local social and cultural influences can also come to shape them. Because of the teacher's role within this environment, as well as the kind of local and specific knowledge that this perspective suggests might help a teacher to improve, interpretive research perspectives are supportive of the idea of teacher research, in which teachers engage in the systematic study of their own practice using some of

the tools and perspectives of interpretive research (Erickson, 1986, p. 156), in addition to others that their unique positions in the classroom and their particular and situated knowledge recommend (see, e.g. Ball & Lampert, 1999; Burnaford, 2001; Hiebert, Gallimore, & Stigler, 2002; Lampert, 1985, 2001; Lee, 2001).[15]

…

Overall, researchers in this tradition are, like process-product researchers, interested in what occurs within the classroom but focus their attention on thought processes and related practices of the teacher and learners. They examine classroom thought or interaction, rejecting what they see as the behaviorism of much of the process-product work that often assumes that the same stimulus will meet with the same response within and between classrooms. This line of work instead examines how individual teachers and students make sense and meaning of the teaching-learning process—given their background, knowledge, experience, or ability, and often in highly variable and situation specific ways. Teacher quality here includes the teachers' ability to recognize these features of the teaching-learning process and to manage them successfully; however, research in this tradition also emphasizes students' role in this relationship and the challenges therein. This view of teacher quality suggests it could be improved by acting on the knowledge bases from which teachers draw in making sense of what and how to engage their students in learning, and how to anticipate or respond to their students' thinking.

IV: Teacher quality as a function of the organizations or systems of which they are a part

Research in the previous category casts teachers and students as thinking beings whose classroom interactions support or constrain students' opportunities to learn. From this

---

[15] Many advocates of teacher research see teacher research as a separate or at least notably distinctive type of research on teaching. Burnaford (2001) explains, "teacher research has different purposes, different incentives, and a different audience than traditional academic research" (p. 50; see also Cochran-Smith & Lytle, 1990).

perspective, instruction can be understood to involve the interdependent interactions among teachers and students around academic content and materials, situated in particular environments over time (see, e.g. Cohen & Ball, 1999; Cohen, Raudenbush, & Ball, 2003; Lampert, 2001). Although studies designed to illuminate cognition, knowledge, and interaction within the classroom do not ignore the potential of the classroom surround to influence these processes, the organizations and systems in which classroom processes are embedded are not an explicit focus of their analyses.

One reason for the lack of attention to forces outside of the classroom may be that the United States' educational system is famously decentralized and uncoordinated, with what happens in the classroom often appearing to have little to do with what happens in other classrooms at that grade level in the school, let alone across the district, state, or nation (see, e.g. Boyd, 1978; Cohen & Spillane, 1992). Weick (1976) and Meyer and Rowan (1978) suggested that this "loose coupling" may be a direct consequence of ambiguities in the instructional process. They theorized that in the absence of a technology of instruction that clearly links what teachers do with what students learn educational organizations attempt to maintain their legitimacy by developing seemingly rational procedures and structures that buffer them from closer inspection that would reveal their decidedly non-rational mechanics and disjointed character. Rowan (1990) noted, however, that more recent research on teaching challenges some of the assumptions of loose coupling theory: "Loose coupling theorists assumed that educational goals are vague, whereas current research on teaching holds that teaching goals are reasonably clear, but dynamic and multiple" (p. 357). While loose coupling theory questions the strength of a knowledge base linking teaching and learning, more recent scholarship on teaching like that of Shulman and his colleagues argues "the work of teachers is complex and nonroutine," but at the

same time, "empirical connections between various teaching strategies and student outcomes are both demonstrable through research and known to expert or master teachers" (p. 357).

Regardless of the cause, this lack of centralization and coherence does not always function to shield teachers—and the schools, districts, and even states in which they are embedded—from their social, cultural, economic, political, and idiosyncratic organizational surrounds. In fact, the fractured character of the American system can leave educators and administrators vulnerable to these forces. In other words, despite the lack of a coherent system through which teaching and learning might be influenced, and perhaps in part because of this absence, the environments in which classrooms are situated can have a powerful influence on the effectiveness of the processes within.

Teachers as actors in organizations and systems

Many scholars have given a great deal of consideration to the formal and informal systems in which teaching and learning are positioned, hypothesizing that an enhanced understanding of the broader context may contribute to better insights about teaching and how it might be improved as an activity and a profession. Cohen, Raudenbush, and Ball (1999) explained why environments might be so important: "Teachers and students shape environments by what they notice and how they respond, but environments shape attention and response" (p. 127). Given this relationship between instruction and its environment, Cohen et al. argued against dividing them into separate categories. They explained, "Teaching and learning are not simply internal technical work that external environments influence, for teacher and learners work, inside instruction, with and on elements of what is conventionally thought to lie beyond practice" (p. 127).

This view of instruction still concludes that the teacher holds a unique place in this system of interactions. Cohen and Ball (1999) note,

> Teachers' knowledge, experience, and skills affect the interaction of students and materials in ways that neither students nor materials can. That is because teachers mediate instruction: their interpretation of educational materials affects curriculum potential and use, and their understanding of students affects students' opportunities to learn. (p. 4)

However, while this view suggests that the *use* teachers and students make of the resources of their environment is what is ultimately important for students' learning, it also draws attention to the actors' *access* to these resources from their educational system and social world—and to the idea that instructional improvement would require attention to both access and use (Cohen et al., 2003). Here, teacher and teaching quality cannot be wholly understood as an individual trait. Instead, for individual teachers and the profession of teaching in the United States, quality is understood to be constituted by complex interactions between the resources to which teachers have access through the organizations and systems in which they are embedded, and their use of these resources. Importantly, these resources can include various types of capital beyond simply "money or the things that money buys" (Cohen, et al., 2003, p. 120). Other critical resources include time, personal resources (e.g. knowledge, skill, and motivation of educators and students), and social resources (e.g. relationships among people, shared norms, and beliefs). This views are in line with recent work on cognition, in which researchers have placed an even stronger emphasis on the social and interactive aspects thinking and learning. From this "situative perspective" cognition is understood to be "(a) situated in particular physical and social contexts; (b) social in nature; and (c) distributed across the individual, other persons, and tools" (Putnam & Borko, 2000, p. 4). The first point suggests that individuals' thinking and learning are inextricably linked with their physical and social contexts. The second posits,

"interactions with the people in one's environment are major determinants of both what is learned and how learning takes place" (p. 5); the various communities in which one participates provide cognitive tools and enculturation experiences that shape thinking. The final point suggests that cognition does not just occur within individuals but is distributed across the individual, other people, and physical and symbolic tools—potentially enhancing the capabilities of each of these elements (Putnam & Borko, 2000).

Empirical investigations of the ways in which organizational or systemic opportunities and constraints come to influence teaching and learning can be conducted in a variety of ways. For example, history, ethnography, intensive interviewing, case studies, or practitioner research can be employed to understand how the institutions and occupations relevant to schooling developed and currently function, as well as how teachers and students make sense of the various competing demands that are made on them within and outside of the classroom walls. Carefully designed programs and their evaluations can illuminate which, if any, of these organizational or systemic influences are amenable to change, how they influence classroom processes, and with what consequences for the educators and learners operating within them. An array of quantitative methods can be employed to attempt parse out the unique influences of these various forces, and their relative strength. Quantitative approaches that take such a systemic view, however, have had to contend with a variety of statistical challenges that characterize multilevel questions, including dependence among individual responses within organizations, and conceptualizations of schooling that allow for variation in the relationships between variables among organizations in the system as well as models of that variation (see, e.g. Bidwell & Kasarda, 1980; Raudenbush & Bryk, 2002). Although the statistical and conceptual roots of multilevel models were developing throughout the 1970s, it was not until the early 1980s that this approach began to

flourish, and it was another decade before statistical computing programs were developed that disseminated these tools widely (Raudenbush & Bryk, 2002).

Schools, the most proximate organization in which classrooms are embedded, have long been understood as both formal and social organizations with a great deal of potential to influence the resources to which teachers and their students have access in their work (see, e.g., Barr & Dreeben, 1983; Bidwell, 1965; Bidwell & Kasarda, 1980; Bryk, Sebring, Allensworth, Luppescu, & Easton, 2010; Lortie, 1975; Waller, 1932).[16] In their well-known review of the literature on "effective schools," Purkey and Smith (1983) highlighted the importance of school organization for classroom processes. They explained, "Only when the school functions to promote the chance of efficient learning being able to take place within the classroom can classroom or teacher specific interventions have much probability of succeeding" (p. 429). The authors noted that many studies have been done to attempt to understand the characteristics of effective schools, including studies of unusually effective or ineffective schools, case studies, and program evaluations. However, many of these studies were plagued with a variety of conceptual and methodological problems, especially with outlier and case studies. Thus, the process of identifying which schools were "effective" was subject to many of the same problems of measurement and consensus that afflicted efforts to identify successful teachers.

---

[16] Interestingly, two recent studies aim to provide additional insight to the extent to which teacher quality may be understood as a stable trait or as a quality that may vary depending on their school context. Both examine the extent to which value-added estimates of individual teacher effectiveness are stable as they move from one school to another, though they reach conflicting results. Using longitudinal data from North Carolina, Jackson (2011) finds evidence that teachers' measured effectiveness increases after a move, supporting the idea that the "match" between teachers and schools is a key factor in teacher quality. Xu, Ozek, and Corritore (2012) recently investigated the assumption that "teacher productivity is portable" across school settings (p. 4), which undergirds policies that aim to "redistribute" effective teachers. Using longitudinal data from Florida and North Carolina, they compare value-added estimates for teachers who moved from one school to the other before and after a transfer. They found that regardless whether or not a teacher moved from a relatively advantaged school to a disadvantaged one, or the reverse (with advantage being characterized by school poverty and academic performance), teacher performance remained stable or improved only slightly; they argue that the small observed changes are most likely due to regression to the mean.

Nonetheless, looking across the literature on effective schools and school culture supportive of student learning, Purkey and Smith (1983) concluded there was reasonable support for the idea that (a) "organizational and structural variables that can be set into place by administrative and bureaucratic means" and (b) "process variables," or those defining "the climate and culture of the school—characteristics that need to grow organically in a school and are not directly susceptible to bureaucratic manipulation" were both characteristic of effective schools and necessary for effective work in the classroom (p. 443). Important organizational and structural variables included school site management, instructional leadership, staff stability, curriculum articulation and organization, school-wide staff development, parental involvement and support, school-wide recognition of academic success, maximized learning time, and, with a nod to higher levels of the system, district support (pp. 443-444). Key process variables included collaborative planning and collegial relationships, a sense of community, clear goals and high expectations commonly shared, and order and discipline (pp. 444-445). Purkey and Smith acknowledged that turning ineffective schools into effective ones was more complicated than schools simply adopting this laundry list of qualities and characteristics, but they hoped that putting the organizational and structural variables into place could create an environment in which effective leaders might work to establish school cultures and consensus that could support student learning in the classroom.

Many of the features that Purkey and Smith (1983) identified as being critical to effective schools are still being investigated by more current programs of research. For example, drawing upon a rich set of data on Chicago's public elementary schools, Bryk, Sebring, Allensworth, Luppescu, and Easton (2010) sought to better understand how educators might better organize schools in order to support student learning and engagement. They found that following the 1988

policies decentralizing governance of Chicago's public schools, those schools that managed to increase student engagement and performance in reading and mathematics shared five interdependent characteristics. These "essential supports" for school improvement include (a) a coherent system of instructional guidance that organizes the content, methods, and related resources teachers might draw upon in their work; (b) the professional capacity of the staff to make good use of these resources; (c) productive parent-community-school ties; (d) a student-centered learning environment; and (e) effective school leadership that helps to focus a school's efforts on cultivating and coordinating elements of the previous supports as well as trusting relationships and collaboration among people within the school. The authors also found that other features of schools, including school size and enrollment stability, and of the broader school context, including the extent of the disadvantage in particular communities, influenced the emergence and efficacy of the essential supports.

In fact, because schools are embedded in other formal and informal systems that have the potential to support or constrain what happens within them, other scholars also expanded their view of the systems relevant to teaching and learning further, hoping that a more systemic approach to studying and intervening in education could give them more powerful tools to reach a greater number of classrooms. Though teacher and teaching quality is often not the stated focus of their studies or the direct object of their interventions, researchers and reformers who take this perspective generally acknowledge that any changes in structure must translate to changes in instructional dynamics within classrooms if they are to influence students' learning. Different scholars aimed to study and leverage different types of systems to this end, however.

In their well-known essay, "Systemic School Reform," Smith and O'Day (1990) proposed reorganizing the formal educational system to catalyze changes in teaching and

learning. Writing at a time when many observers were clamoring for more academically demanding and ambitious instruction, as well as evidence of this learning through improved student outcomes (see, e.g., Cohen, Moffitt, & Goldin, 2007; Mintrop, 2004), Smith and O'Day recognized that these changes would require most teachers to transform the way they taught. They explained that for precisely this reason, reform efforts must be systemic rather than taking place in on a school-by-school basis. According to Smith and O'Day's formulation, the states could take on a critical role in this process. "States not only have the constitutional responsibility for the education of youth," they explained, "but they are the only level of the system that can influence all parts of the K-12 system: the curriculum and curriculum materials, teacher training and licensure, assessment and accountability" (p. 246).

Smith and O'Day (1990) proposed that by holding schools accountable for student achievement, measured by assessments aligned with challenging, state-designed curriculum frameworks, other important components of the states' educational systems would be forced into alignment as well. This in turn would create a coherent system of instructional guidance at the state level. In addition to creating more consistent systems for teachers, students too would experience greater coherence as they moved through the grade levels of the system. Though the state departments of education would be primarily responsible for implementing standards-based reform, in Smith and O'Day's vision this centralized system would be balanced by a degree of local flexibility and professional prudence. Local schools and teachers would be left to determine how to best prepare their students to meet the objectives set by the state. Smith and O'Day emphasized that upon implementing such reforms, many teachers would not have the capability successfully carry out this critical piece—designing and enacting quality instruction aligned with state standards—and that districts and states would have difficulties in supporting them due to

their own challenges with competence. However, they imagined that a coherent system of instructional guidance in which components of the educational system such as teacher certification and professional development were aligned would provide teachers with the chance to learn how to do them. The theory of reform Smith and O'Day provided can be understood as roughly that which is instantiated on a national scale by the No Child Left Behind act of 2002, though in practice their ideas have not always worked out as they intended.[17]

Despite the powerful influence of the idea of systemic, or standards-based reform organized at the state level, there are other ways of conceptualizing of educational systems in which schools and classrooms can be embedded and through which instructional quality might be enhanced. For example, beginning in the 1990s, a variety of multi-faceted school reform programs, or comprehensive school reform (CSR) programs, were designed and implemented by scholars and educators who saw a great deal of promise in whole-school approaches to instructional improvement (Rowan, Camburn, & Barnes, 2004). These CSRs are most frequently run by external non-profit or for-profit agencies, and reach across traditional governance hierarchies to engage schools in multiple districts and states. CSR programs have remarkably different organizational strategies for supporting changes in school management and instruction (Rowan & Miller, 2007). They also vary in their aims for school management and instruction, as well as in their effectiveness in implementing their reforms in ways that change teachers' practice (Correnti & Rowan, 2007) and contribute to students' achievement gains (Borman,

---

[17] Carnoy and Loeb (2002) conducted a study suggestive of the possible utility of standards-based reform. They examined the relationship between stronger centralized accountability mechanisms and students' mathematics achievement on the National Assessment of Educational Progress (NAEP) from 1996-2000. They found a "positive and significant relationship between the strength of states' accountability systems and math achievement gains at the 8th grade level across racial and ethnic groups" (p. 320), and that "states with stronger accountability saw significantly greater gains in the percent of 4th grade Black students that achieved at least the basic level on the math NAEP […]; and marginally significant greater gains in the percent of 4th grade Hispanic students that achieved at least the basic level on the math NAEP" (p. 321). However, there remained considerable variation among student outcomes "among states with similarly weak or strong accountability systems" (p. 321), suggesting strong accountability systems are not necessarily linked to better and more equitable outcomes.

Hewes, Overman, & Brown, 2003). However, they can be understood as another systemic effort to improve teaching and learning across multiple school settings by attempting to create and manage their own coherent systems of instructional guidance and support.

A related set of efforts has focused on developing the profession of teaching and the knowledge base from which teachers can draw. The promise of this approach is that a shared system of professional values and expertise instantiated in shared knowledge, practice, and tools could support teachers in learning and making sense of their work—especially the particular and nonroutine aspects of it—and for professionally responsible adaptations of these resources in the contexts of their work (Hiebert, Gallimore, & Stigler, 2002, Weick & McDaniel, 1989). In this vein, there is some evidence that professional communities and collaborative groups operating within or across schools can provide teachers with knowledge, social supports, and opportunities to learn that allows them to change their practice in ways they believe to be more effective (see, e.g., Lieberman, 2000; McLaughlin & Talbert, 1993). Scholars including Bryk (2009) and Raudenbush (2009) argue that such efforts are more likely to support reliably strong instruction if professional learning communities are organized around a shared "instructional regime," or "instructional system" that would stipulate what teachers need to know about their students and their "background knowledge, skills, and interests," as well as include "very specific pedagogical practices and social routines and expects automaticity in their use" (Bryk, 2009, p. 600). Working together within this system would allow educators to access shared notions of what constitutes student learning, as well as a common language with which to their discuss progress toward these goals. Bryk offers Montessori education and Reading Recovery as examples of "the organizing power of a specific instructional system on the activity of a professional learning community" (p. 600). These collective approaches could take place within the traditional

contexts of schools or school districts, though this would go against cultural norms of teaching within the United States (Cohen & Ball, 1999; Elmore, 1996; Lortie, 1975; Raudenbush, 2009).

Other researchers have attended to a different set of systems—those broader social, cultural, and economic systems that influence how students and teachers approach what occurs in classrooms. For example, Powell (1996) pointed out that American teachers are faced with singular challenges regarding the motivation of students. He considers motivation to learn to be an American dilemma because "American society has made an unprecedented commitment to mass schooling, but it provides weak incentives to learn and relatively strong disincentives not to" (p. 21). With compulsory universal schooling in a decentralized system, students neither choose to be in school nor have to compete to maintain their positions there. What is more, there is no widespread consensus about what students should know, common incentives for them to do so, or consequences if they do not. Coupled with the "yawning gap between school and business" (p. 30) and "the tradition of American anti-intellectualism" (p. 31), it begins to seem irrational for students to bother working particularly hard in school. Powell noted that this problem of motivation has largely been delegated to individual teachers who are dependent upon their students' engagement for their own success. Cohen (2005) observed that these social, cultural, and institutional arrangements, among others, vary within and across societies and create drastically different conditions for teachers' practice. He explained, "These arrangements also increase the probability that extraordinary expertise will be required to get barely decent results in some situations, while making it possible to get fine results with only modest expertise in others" (p. 288). This again points to the idea that the ability to draw positive results from groups of students is not a fixed asset of the teacher, or even of the interactions occurring within the classroom, though that is where it is activated.

How to manipulate these systems in ways that create conditions supportive of improved teaching practice and student learning, however, remains somewhat unclear. What is more, interceding in some of the social, cultural, and economic systems in which schools and classes are embedded are gargantuan tasks that would require overstepping the boundaries of traditional education policy or finding ways to operate outside of them. Additionally, the evaluation of any systemic variations or intentional interventions, as they relate to teacher or student learning, are subject to the same difficulties that are endemic to types of research on teacher and teaching quality, including questions of the adequacy of the assessment instruments and their alignment with the curriculum, pedagogy and goals of the teacher and the rest of the system in which he or she is situated.

…

In summary, researchers in this tradition focus on teachers as members of organizations or systems that support or constrain their work. Teacher and teaching quality are understood to exist in the dynamic relationships among students, teachers, content, and the surrounding social, cultural, and material contexts that come to bear on these relationships—and ultimately on learning. Attention is given to the coherence of instruction across time, classrooms, or parts of the organization—especially as teachers and students experience it—and to the resources that teachers and students are able to draw upon in their work together. In addition to involving a more complex conceptual framework for teacher and teaching quality, this line of research requires complicated methodological tools that are sensitive to these various influences. From this perspective, improving teacher quality involves acting on the schools, school systems, and social systems of which teachers and students are a part in addition to teachers' knowledge and skill in making use of these resources. Relevant observations about the dynamic relationship

between schooling and its organizational and social surround have long been part of the discourse about U.S. education, but contemporary research methods, as well as recent interventions through policy and private organizations have encouraged the prominence of this perspective on teacher and teaching quality in the past several decades.

V: Teacher quality as a function of student achievement

Adjusted student achievement scores have long been used as a criterion by researchers interested in the empirical study of teacher quality (see, e.g., Crabbs, 1925; Walker, 1935). However, researchers working in the tradition described by this final category do not attempt to link these gains with teachers' traits, behaviors, knowledge and reasoning, or contexts. Teacher quality is *defined* simply as a function of a teacher's contribution to the gains his or her students make in their academic achievement. In other words, good teachers are those who promote student gains in achievement. From this perspective, since teachers cannot be effectively sorted by who they are or what they do, and because we cannot determine what they should do to be more effective, they should be assessed only by the outputs of their classroom. Assessing teachers on the size of their students' achievement gains, offering incentives as rewards and motivation for average or high performers, and removing those who do not make the grade then appears to be the most logical way to improve teacher quality. This approach may seem to be the most parsimonious way to measure teacher and teaching quality since it focuses directly on the outcomes that so many researchers have understood to be the most important criterion of teacher quality. However, while this line of research has helped to establish the deep importance of teachers to student learning, it also paradoxically renders teachers and teaching somewhat invisible.

Teachers' value-added

In recent decades, efforts to isolate and quantify the extent of teachers' influence on student learning have involved a group of statistical techniques that are often referred to as "value-added models," or VAMs (see McCaffrey, Lockwood, Koretz, & Hamilton, 2003 and Lipscomb, Teh, Gill, Chiang, & Owens, 2010 for reviews). Braun (2005) offers a useful summary, "Essentially, VAMs combine statistically adjusted test score gains achieved by a teacher's students. Teachers are then compared to other teachers in the district based on these adjusted aggregate gains" (p. 5). Despite these similarities, "Various VAMs differ in the number of years of data they employ, the kinds of adjustments they make, how they handle missing data, and so on" (p. 5).

Widespread enthusiasm about VAMs and their potential is motivated by familiar hopes about the solution of practical problems in education. In fact, use of VAMs has helped researchers to make some important contributions to the scholarship on teacher and teaching quality. Perhaps most importantly, researchers employing various types of VAMs have established a solid empirical record demonstrating that two fundamental assumptions underlying research on teacher and teaching quality seem to be true—first, that teachers vary in the extent of their influence on their students' learning, and second, that these differences can be large and consequential (e.g., Chetty, Freidman, & Rockoff, 2011; Rivkin, Hanushek, & Kain, 2005; Sanders & Rivers, 1996; Wright, Horn, & Sanders, 1997).[18]

Yet VAMs seem promising as a tool for research and policy for other reasons. For example, they appear to offer a different approach to "accountability" than that offered in traditional standards based reform legislation, and NCLB specifically. NCLB requires that schools demonstrate adequate yearly progress (AYP) towards a common achievement target for

---

[18] Taking advantage of the random assignment of students and teachers to classes that was part of the Tennessee Class Size Experiment, or Project STAR (Student-Teacher Achievement Ratio), Nye, Konstantopoulos, and Hedges (2004) estimate within school teacher effects that are similar to those calculated using sophisticated VAMs.

successive cohorts of students by grade level and subgroup. VAMs seem to provide a way of measuring the contributions that teachers (and schools) make to individual students' progress over time (Braun, 2005; McCaffrey, et al., 2003). In addition, VAMs are used in many states and school districts around the country because administrators want to identify individual teachers who might be candidates for termination, targeted professional development efforts, promotions, or salary increases. VAMs are also likely to play an increasingly important role in the evaluation of teacher education programs and a variety of other education programs and policies (Braun, 2005; Gansle, Noell, & Burns, 2012).[19]

Limitations to value-added modeling

Despite the enthusiasm VAMs have generated among policymakers and educators, a variety of scholars have urged educators, policymakers, and other researchers to use caution when interpreting and applying their results or making them public due to a host of technical and theoretical considerations, many of which echo the difficulties encountered by others in attempting to use student achievement gains as criteria of effectiveness (see, e.g. Braun, 2005; Kupermintz, 2003; Hanushek & Rivkin, 2006; McCaffrey, et al., 2003; McCaffrey et al., 2009; Rothstein, 2008; Sass, 2008). Indeed, a great deal of concern remains about the extent to which VAMs are able to isolate the causal influence of a teacher on his or her students' learning, and compare such effects fairly. This concern arises because of the myriad of influences that come to bear on the teaching and learning process, which are of primary interest to other researchers of teacher and teaching quality. For example, neither students nor teachers are randomly assigned to classrooms, and there may be patterned differences in their distribution that remain unaccounted

---

[19] Many of these states and districts use VAMs that make somewhat different assumptions about the measurement of teachers' effectiveness (see Braun, 2005 and Lipscomb et al., 2010 for overviews of several of these systems how they differ). Researchers also sometimes develop their own VAMs for the purposes of their studies.

for in some VA estimates. In addition, school or district resources and policies that are not within a teacher's control may be unevenly distributed and may influence the degree to which they are able to support their students' learning. Hanushek and Rivkin (2006) explain that using VAMs, which include controls for prior achievement, assuages "… problems resulting from the lack of historical information, [but] it does not protect against the confounding influences of contemporaneous factors related to the variables of interest and not captured by prior achievement" (p. 1060). Ignoring the potential bias introduced by these issues can lead to what Braun (2005) refers to as "inappropriate attribution," through which "teacher can be inappropriately credited or penalized for their students' results" (p. 8).

While it certainly can be argued that teachers' "quality" resides in their abilities to support their students regardless of the circumstances, this approach to measuring teacher quality can be problematic when consequential decisions, such as job termination or large salary changes are solely contingent upon teachers' estimated effectiveness. For this reason, some researchers have argued in favor of examining only within-school variation in teacher quality, "eliminating both the actual variance in teacher quality between schools and any observed student, community, and school differences including the impacts of principals and other administrators" (Hanushek & Rivkin, 2006, p. 1067). However, this tactic introduces other possible problems due to smaller numbers of teachers and volatility due to staff mobility, and raises ethical questions about the peers against which teachers' relative effectiveness ought to be evaluated.

In fact, each of the models carries assumptions that, if violated, can result in poor estimates of teachers' effects. Braun (2005) noted that many models assume that within grade levels, all teachers have been assigned similar academic goals and have access to similar resources—two conditions prior research on American schools suggest is quite unlikely. Within schools, tracking

may differentiate the goals of the classes, and across schools within districts, differences in curricula are common. In addition, the precision of the estimates provided by VAMs depend, in part, on the models themselves and the amount of data they use. For example, some opt not to include student background characteristics, which appear to be somewhat weakly related to changes in educational achievement (Braun, 2005, p. 10). Questions about the stability of the estimates year to year, given the small numbers of students teachers sometimes have in their classrooms and missing data, have led some analysts to average several years of data for estimates of effectiveness, while others use a single year (Braun, 2005; Goldhaber & Hansen, 2010; Hanushek & Rivkin, 2006; McCaffrey et al., 2009).

Familiar concerns with the tests themselves, including their validity and the ways in which they are scaled, are also prominent in debates about VAMs and the appropriate application of their estimates. Unsurprisingly, VAM estimates are sensitive to the content of the tests themselves, when they are administered, the measurement error associated with them, and whether or not they are high- or low-stakes exams for teachers (see e.g., Corcoran, Jennings, & Beveridge, 2011; Papay, 2011; Sass, 2008). Many of the test used for creating VA scores are not designed for these purposes. For example, Hanushek and Rivkin (2006) write,

> [M]ost achievement tests are not designed to provide valid rankings of the effectiveness of teachers with very different mixes of students in terms of academic preparation. For example, a test that concentrates on rudimentary material will do a poor job identifying differences in teacher quality among teachers whose students could answer the vast majority of questions on the basis of knowledge acquired prior to the current school year. (p. 1066)

Horizontal scaling, in which different forms of the same test within a grade are equated, and vertical scaling, in which different test instruments across grades are placed on a single scale, both introduce uncertainty into the measurement of achievement growth, and, particularly in the upper grades, may conceal information about the types of progress students are making (Braun,

2005). Scaling across grades also brings up other questions. Braun (2005) explains, "Aside from the technical aspects of vertical scaling, there is a question of what it meant to put, say, third-grade and seventh grade mathematics scores on the same scale. In particular, should we treat a 20-point gain at the low end of the scale as equivalent to a 20-point gain at the upper?" (p. 14). In addition, the alignment of the tests to what they purport to measure—which, with NCLB has largely become the state standards—is questionable, especially when state standards aim to measure knowledge and skills that are not easily captured in typical multiple-choice questions.

Another set of very basic questions about the utility and fairness about the application of VAMs have to do with the extent to which current state and district data systems can accurately connect student tests scores with the teachers who were responsible for their instruction in a particular time period or subject areas. While this may seem relatively easy, a close look at the ways that students and teachers are organized in schools reveal that this is not the case. For example, student mobility as a result of family circumstances means that students may move in and out of schools and classrooms throughout the year, something that is particularly common in high-poverty schools; students are also often regrouped across classrooms and teachers for instruction in different subject areas, sometimes several times a year. Currently, the vast majority state and district data systems are not designed to capture this complexity, meaning that research or administrative decisions based on their results may be inaccurate (Batelle for Kids, 2009).

Given concerns of this kind, many have concluded that VAMs are not yet sufficiently stable or reliable tools to use for making consequential decisions about a teacher's effectiveness and his or her livelihood (see, e.g. Braun, 2005, Kupermintz, 2003; McCaffrey et al., 2003). Even Hanushek and Rivkin (2006), staunch advocates of loosening the entry requirements for teaching and then improving teacher and teaching quality via incentive structures based on some

67

variation of VAMs, caution that schools may want to balance single measures of teachers'

"quality" with "evaluations of overall effectiveness" by administrators who theoretically could

capture "more comprehensive and nuanced" information about teachers' performance (p.

1072)—perhaps like the teacher rating scales that have long been a feature of schemes to

evaluate American teacher quality (see also the discussion of the findings of the MET project in

the next section of this essay). Despite such concerns, Braun (2005) recommends that VAMs be

used as tools to identify teachers who are likely to need professional development, and for

identifying underperforming schools.

Besides possessing a series of technical challenges, VAMs are limited in other important

ways. In her 2005 AERA Presidential Address, Marilyn Cochran-Smith pointed out that using

test score gains to define teacher quality leaves much left undiscovered. She explained,

> At the end of the day, then, teacher quality remains a black box—we do not know
> what effective teachers do, know, believe, or build on, nor do we know the
> conditions that make this possible. Further, because teacher quality is isomorphic
> with pupil achievement—achievement, too, is a black box, and we know nothing
> about what and how high performing pupils learn, what resources they bring to
> school with them, or how they build on what they know. (pp. 6-7)

On its own, this approach to defining and measuring teacher quality could provide important

feedback to teachers about their work, but offers little direction for teachers who wish to improve

their practice or for others who would like to design professional development and training,

curricula, or other forms of instructional support.

Another non-technical concern about VAMs is related to what they suggest our society

expects from our schools. Rivers and Sanders (2002) explain, "To some, equity in educational

delivery will be achieved only when simple group averages across various demographic

subgroups are equal" (p. 14). They contrast this "simplistic" definition of equity with a more

"realistic" one: "If true equity is defined as each student making appropriate academic growth

each year, then expectations for *educators and students* can be set in terms of academic growth rates" (p. 14). They reason, "If appropriate rates of academic growth are sustained across grades, then *all* students' academic attainment will be ratcheted to higher levels" (p. 14). However, given the fact that young children enter school differentially prepared for school, and that poor and minority students in particular begin school, on average, well behind their white and more affluent peers (Lee & Burkam, 2002), an approach that focuses solely on gains may not be all that we choose to ask of our teachers, schools, or school systems.

…

The fifth and final line of work I identified in this essay defines teacher quality solely on the basis of student achievement outcomes and focuses on measuring teacher effects without speculating on the underlying causes. Methodologically, isolating these effects are challenging for a host of reasons. While this approach can be understood as the one measuring teacher quality most directly, it can also be seen as ignoring teachers and teaching—who they are, what they do in the classroom, why, and how they might improve the ways they work together. This research has played an important role in establishing that teachers can have an have a consequential impact on students' learning trajectories, but without joining this work with those of the other traditions to link students' learning with teachers' characteristics, qualifications, behaviors, or knowledge, or with characteristics of the school system—this approach does little to guide policymakers and practitioners in ways to improve the quality of the teaching force.

Working across traditions

In part because of research using VAMs that has identified teachers as the most powerful school-based influence on student learning, in recent years there has been renewed interest and activity

around developing measures of teacher and teaching quality that may be used for professional development, teacher evaluation, or both. For example, the U.S. Department of Education's (2010) Race to the Top program required states competing for grant money to agree to:

> Design and implement rigorous, transparent, and fair evaluation systems for teacher sand principals that (a) differentiate effectiveness using multiple rating categories that take into account data on student growth (as defined in this notice) as a significant factor, and (b) are designed and developed with teacher and principal involvement. (p. 34)

The largest and most prominent example of this approach in the research community is the Gates Foundation's $45 million dollar investment in the Measures of Effective Teaching (MET) project, launched in the fall 2009 with the aim of identifying teaching practices associated with student learning gains. The foundation set out to design or validate tools that states and school districts might use to evaluate teachers and to provide them with feedback for their improvement, arguing these resources would provide a new "knowledge base for practitioners and policymakers who are trying to strengthen the teaching profession" (Gates Foundation, 2012).

More specifically, the MET project was designed to enable researchers to investigate the relationship between student outcomes—including value-added estimates using state assessments as well as more ambitious supplemental assessments, and student reports of effort and enjoyment—and four types of process measures. These process measures included (1) measures of teachers' pedagogical content knowledge; (2) surveys of student experiences in the classroom; (3) surveys of teachers about school working conditions, including instructional support; and (4) scores from multiple instruments for assessing videotaped classroom observations, the reliability of which researchers also sought to determine as they were applied at a relatively large scale. Some of these observation protocols were designed for use across subjects, including the Classroom Assessment Scoring System (CLASS) developed by Robert Pianta and the

Framework for Teaching (FFT) developed by Charlotte Danielson. Others are subject specific, including the Mathematical Quality of Instruction (MQI) protocol developed by Heather Hill and Deborah Ball; the Protocol for Language Arts Teaching Observations (PLATO) developed by Pam Grossman; and the UTeach Teacher Observation Protocol (UTOP) developed by Michael Marder and Candace Walkington for use in science and mathematics, but used by the MET project for assessment of mathematics instruction.

Discussing his foundation's investment in the project with a New York Times columnist, Gates explained that he had long been interested in questions about teacher effectiveness such as "How do you measure it? What are the skills that make a teacher great?" Gates is quoted as remarking, "It was mind-blowing how little it had been studied" (c.f. Nocera, 2012). However, the MET project, and the research of several of the scholars affiliated with it, is a clear descendent of each of the five traditions presented earlier. For example, the MET project's primary focus is developing provide insight into teaching practices that influence student learning, and in this way may be understood as a descendent of the process-product tradition of research. Nonetheless, many of the instruments that are part of the study are more involved than either the high or low inference observation protocols typically used to describe teacher "behaviors" in early process-product research, and draw carefully from the insights of research on teaching that emphasize teaching quality as a function of social interaction, cognitive processes, and interaction among teachers and students that shape and are shaped by their conceptions of the subject matter being studied and the broader contexts of their work.[20] While

---

[20] For example, the CLASS system developed by Pianta and his colleagues embraces a view of classrooms as social organizations. CLASS uses 7-point rating scales to assess the quality of the classroom environment on a variety of dimensions (Hamre & Pianta, 2007). The three overarching dimensions include the emotional supports, classroom organization, and instructional supports observed in the classroom, each of which its authors assert is linked to a strong research base suggesting its importance for student learning. Within these broad dimensions, a variety of items that vary somewhat across grade levels are used to assess more specific features of the classroom climate and interactions, including, for example, a teacher's sensitivity and regard for student perspectives, his or her behavior

challenging the use of VAMs alone for assessing effectiveness, the MET study draws upon the methodological tools developed by researchers working in (or challenging) that tradition.

In January of 2012, Gates released their most recent report based on 7,491 lessons from 1,333 teachers in six school districts from around the country who taught mathematics or English language arts (ELA) to students in grades 4 through 8. Survey data from over 44,500 students was also used in their analyses. Videotaped lesson observations were each scored at least three times using the CLASS, FFT, and the appropriate subject-specific instrument by observers who had received substantial training. The authors highlighted five key findings that generally supported their approach. First, they found that each of the classroom observations instruments tested was associated with student achievement gains regardless of the assessment used. Second, they discovered that since "a teacher's score varied considerably from lesson to lesson as well as observer to observer" (p. 8), obtaining a reliable observation score required them to score four separate lessons by different observers. Third, they found that adding student feedback and a measure of a given teacher's achievement gains with a separate group of students greatly improved the predictive power of the observation protocols, and that combining the three also improved the measures' reliability. While observation protocols did little to add to the predictive power of value-added estimates, and are less reliable than feedback from student surveys, their use in conjunction with the value-added information also provides information about teachers'

---

management and productivity, and the quality of feedback and language modeling offered by the teacher. Despite a notably different approach to investigating classroom processes, Pianta and his colleagues share a focus with interpretive researchers on the classroom as a social system in which students' learning opportunities are mediated through their interactions with teachers. Meanwhile, PLATO—at least in its original pilot version described in Grossman et al., 2010)—was designed as a supplement to CLASS for middle and high school English language arts (ELA) classrooms. Therefore, Grossman and her colleagues adopted the CLASS 7-point scale and developed indicators referenced to this scale used to score ten elements of instruction found to be "effective" in previous research—including things like "clarity of purpose of the lesson," "level of intellectual challenge in both teacher questions and tasks assigned to students," "use of models and modeling of both high quality work and strategies for reading and writing," "presence of explicit strategy instruction in reading and writing" and "accommodations for English Learners"—as they are articulated in secondary ELA classrooms. PLATO also captures information about the focus and content of instruction via a checklist of major content domains.

practice that can be used to support their improvement, something that is not provided by the value-added scores or student survey results. Fourth, they found that the combined measures were more predictive of student achievement gains then either teacher experience or holding a master's degree. Finally, teachers who scored highly on the combined measure also had students who made greater gains in assessments of conceptual understanding in mathematics or short-answer literacy assessments, and who reported greater effort and enjoyment in their classes (for more support for the use of a combination of subjective observations and VAMs to evaluate teacher effectiveness, see Rockoff and Speroni, 2010).

While these and the more nuanced findings of the MET project—and those of the scholars affiliated with the project or its instruments, like Grossman et al., (2010) and Pianta (Hamre & Pianta, 2007)—make useful and worthwhile contributions to our understanding of instruments and practices that are associated with student learning gains, these recent studies share the optimism and general approaches of previous generations of researchers on teacher and teaching quality, but also, as I explain below, some of their challenges.

More and better research is not enough: The American educational context

Looking across these five traditions of research, and recent studies and instruments that draw from them, it is apparent that each has contributed something to our understanding of teacher and teaching quality, with some hints about how teaching might be improved as a practice and profession. Together researchers have learned, for example, that teachers' characteristics and qualifications are generally insufficient for differentiating among more and less effective teachers. However, it seems teachers do tend to become more effective in their first several years of teaching, and that verbal ability, academic ability, and content knowledge are important if not

sufficient resources for teachers' effectiveness, particularly in some subject areas and grade levels. Scholars have confirmed that what teachers *do* in classrooms matters for their students' learning, but also that students make sense of this instruction in variable ways depending on who they are and their particular social environments. The way that teachers understand and make use of the various forms of instructional guidance with which they are bombarded, of the subject matter, of their role, and of their students informs their practice, which in turn influences the ways in which students are positioned to learn. Broader social, cultural, and material resources shape what teachers and students are able bring to their work, and how they interact with each other and the content to be learned. Whether and how teaching and learning is measured, and the extent to which teachers, students, and the public find meaning in these assessments also matters, and are topics that continue to be the subjects of important and lively debates in research, policy, and practice.

Despite these important contributions, there remains a great deal of controversy about what constitutes teacher and teaching quality, and therefore how to work towards improvement in these areas (see, e.g., Cochran-Smith, 2005; Gates Foundation, 2010; Goe, 2007b; Hanushek & Rivkin, 2006; Hess, Rotherham, & Walsh 2004; Jerald, 2012; Rice, 2003). There are several ways to understand why these numerous and diverse efforts have not resulted in greater consensus about teacher and teaching quality. One view is that technical difficulties have impeded progress in the development of a knowledge base about teacher quality. From this perspective, as research instruments, statistical techniques, and the availability of data and data systems linking student outcomes with teachers continue to improve, studies that clearly link teachers' attributes, their practices, and their students' learning should emerge. In the early decades of the twentieth century, this idea was quite widespread in part because the field of

measurement of social and cognitive phenomenon was still in its early years and seemed to hold so much promise. This perspective remains popular, however, and not without reason given advances in this area over the past century, including models that attend to the hierarchical or nested nature of educational data, various VAMs, and increasingly sophisticated, if still limited, data systems (see, e.g. Goe, 2007b; McCaffrey et al., 2003).

A second possibility for this fragmentation of accord with respect to teacher and teaching quality is that there has been a belated focus on instruction and student learning outcomes. From this standpoint, as researchers turn their attention away from a simple focus on "inputs" and instead focus on the processes by which they are used by teachers and students, and with what results, a greater consensus on teacher quality will emerge. This seems reasonable as well because it was only in the 1960s that instruction and student outcomes became central to studies of educational quality, and the theories, instruments, and methods on which they are based have developed considerably during the intervening decades. Research examining teacher and student action, knowledge, thinking, and interaction has indeed contributed to our conversations about the teaching-learning process and how it might modeled, measured, or improved.

Yet a third prospect is that the educational process is too complex and contextual, and so reliable relationships between measures of student learning and teacher characteristics, knowledge, and practices can never be established, let alone legislated. This view suggests that additional research trying to link these pieces of the educational process is likely to remain inconclusive or contradictory. This perspective has had a variety of adherents over time. Members of AERA's mid-century Committee on the Criteria of Teacher Effectiveness (CCTE) (1952) worried about this possibility at the same time as they decried the "meager scientific foundation" upon which systems for teacher selection, training, and supervision were built (p.

238). And both Erickson and Hanushek—in many ways at opposite ends of the spectrum with respect to the study of teacher quality—agree that educational processes may be too multifaceted to support reliable relationships between inputs and outputs into the system. Hanushek (2002) wrote, "We do not know how to identify a well-defined set of inputs that is either necessary or sufficient for ensuring high-quality schooling." He continued, "I believe that the educational process is much too complicated for us to uncover a small set of criteria that are amenable to central legislation and control" (p. 7). Therefore, in his view, it makes more sense to take the approach outlined above in the fifth category of research on teacher and teaching quality, in which incentives for teachers and other school officials are linked with their value-added to students' learning and teachers are fired when their scores place them in the lowest 5-10% of their peers. For Erickson (1986), on the other hand, the complex and contextual nature of the teaching-learning process suggests a method of study that attends to the particulars of practice and the "local" meanings people assign them.

While there is likely some truth in each of these three explanations, I propose that none of them fully explains the fragmentation, or underdevelopment, of an empirical base that establishes what teacher and teaching quality is and how it might be improved. Instead, I argue that a close examination of the research in this area suggests that two of the most *fundamental features of the U.S. educational landscape* undermine the success of these efforts. These fundamental features are (a) the absence of an "infrastructure" for education (Cohen, 2011) that would serve to support and direct teachers' practice; and (b) a political system that maximizes variability in practice, reduces coherence among parts of the system, and multiplies the uncertainties of weak technologies and numerous and conflicting goals for education.

Cohen (2011) introduces the concept of an "infrastructure of practice" (p. 54) to refer to a set of social resources essential for those engaged in what he calls the "practice of human improvement." Cohen explains,

> Skilled occupations could not thrive without the extensive technical and professional affordances that enable work. Those include socially accepted views of the proper terrain of work; the sorts of problems that practitioners can claim to solve and the results that should be expected; the specialized knowledge and skills required to solve those problems; the specialized terminology needed to discuss, plan, perform, and assess work; the education that prepares people to work and to improve; norms and standards that inform judgments about the quality of work; the organized intelligence that influences the invention of new tools and technologies; and procedures to deal with unacceptable work. (p. 56)

Cohen suggests that in education, such an infrastructure would include a shared set of exams to which detailed curricula or curricular frameworks could be referenced, and around which teacher education programs might be designed; a specialized vocabulary that could be used to discuss teachers' practice and student work, and how they might be improved; and a set of professional standards or norms that, together with the exams and curriculum, guide and support teachers in making decisions about what and how to teach. The specialized knowledge and skill of educators would inhere in each of these elements of an infrastructure of practice. Depending on the strength of the infrastructure's design and how well it is used, it might be understood as a way to organize collective and individual professional learning into a dynamic professional knowledge base for teaching, and for learning to teach.

However, in the U.S. system of education, a unified infrastructure of this kind is conspicuously lacking. Its absence, or underdevelopment, is related to the point (b) above with respect to the design of the political system in which American education is embedded. Cohen and Spillane (1992) observed, "the U.S. political system was specifically designed to frustrate central power" (p. 5), and this intention was certainly realized in education. They continued,

"Authority in education was divided among state, local, and federal governments by an elaborate federal system, and it was divided within governments by the separation of powers" (p. 5). State governments, by law the central authority for education, until recently delegated much of their power to localities.

A weak central authority has allowed a variety of governmental and nongovernmental agencies to assert their influence on individuals working at various levels of the educational system (Cohen & Ball, 1999; Kirst & Walker, 1971; Rowan, 2002). Schools and teachers are often caught in the crossfire of local politics, or between the conflicting pressures of students and their families. State standard and assessment programs; private textbook or testing companies; and teacher education and professional development programs—all also subject to political conflict and some of its strengths and liabilities—attempt to influence teachers' instruction. Professional organizations, which often serve as a unifying force in other practices of human improvement, also have attempted to exert some influence on educators. However, Cohen (2011) observes, "[P]ublic school teaching is a wholly owned subsidiary of state and local government. The conditions of teachers' work, their salaries, and standards of quality work are set by state and local governments, not by independent professional organizations" (p. 62). Rarely are these various influences aligned, and often they make competing demands on teachers who in turn take up these influences in often highly variable ways based on their personal capabilities and beliefs (Brophy, 1982; Freeman & Porter, 1989).

Within such a patchy and decentralized system, it is unsurprising that the key elements of an educational infrastructure failed to materialize, even in light of recent reform efforts. Standards-based reform, instantiated in national policy with NCLB, attempts to remedy some of this incoherence in the system by holding schools accountable for their students' learning on

content standards measured by (theoretically) aligned standardized assessments. However, NCLB still delegates the responsibility for these matters to 50 separate states. More importantly, a system of standards and assessments leaves educators without many other crucial elements of an infrastructure of practice that would support them in their work—including a more developed curriculum and program of education that would help them to understand the standards and how their students might best learn them, or professional standards or values that could also help them to manage the uncertainty of their work (Cohen, 1995). Therefore, while standards-based reform offers the promise of greater centralization in governance, many fissures remain and the extent to which these reforms can create greater coherence in practice is questionable.

The decentralized design of the U.S. political system together with a deeply fractured and underdeveloped infrastructure for educational practice has helped to create unrealistic expectations for what research on teachers and teaching quality might accomplish while simultaneously frustrating efforts to study and to apply findings to teaching. The U.S.' strong tradition of local control, powerful disagreements about the aims of education, and the absence of a centralized authority in education—either a governmental body or formally sanctioned professional organization—to establish a set of common objectives for education or agreed upon ways to assess progress towards them have together led many to pin their hopes on their "scientific" derivation. Yet as researchers across the last century have observed, the aims of education and the methods we use to assess progress towards them cannot be determined entirely by science but must involve political and ethical deliberations and decisions that have traditionally been delegated to states, localities, and individual educators who have taken up these considerations with variable care and outcomes.

Even among researchers who agree that student learning of academic content is a desirable outcome of formal education, this educational context continues to cause many problems for the study of teacher and teaching quality and for the integration and application of this work. First, the proliferation and use a wide range of tests made by a wide range of governmental and nongovernmental agencies means that the test results, and those of the studies using them, are not directly comparable. Second, as Shavelson et al. (1986) and others have pointed out, many researchers used, and continue to use, norm referenced tests in their studies of teacher and teaching quality. Norm referenced tests were developed to measure students' achievement relative to one another regardless of the specific curriculum in use in their classroom and were not designed to be sensitive to changes in teachers' instruction. Using results of norm-referenced tests to make estimates about teachers or their methods is therefore potentially quite problematic. Third, lacking a common curriculum or even shared goals among teachers, separating out the various influences of teacher and student background, beliefs, knowledge, content coverage, and practices employed when using these tests as the outcome of interest is likely to be an inherently muddled endeavor (Shavelson, et al., 1986). It also makes it far more likely that rather generic aspects of teacher and teaching quality, such as "time on task" will be found to have a significant influence on student learning than more complex measures of things like pedagogical content knowledge, because the content that teachers are expected to know and cover, and that which they do, and that which is measured, vary so much that their apparent mean impact on learning may be diminished regardless of their actual importance.

The absence of two key pieces of an educational infrastructure—a common set of exams aimed at measuring a shared set of goals, and a common curriculum or curricular frame directed at supporting teachers and students in reaching these ends—has denied researchers not only the

possibility of better studies but also an organized framework around which they might coordinate their efforts and integrate and disseminate their findings. Lacking this frame, researchers have faced difficulties when they attempt to collate and generalize their findings in ways that might be useful for the construction of other aspects of an infrastructure of practice, including a program of teacher education. For example, Charters and Waples (1929), who carefully examined teacher traits and over 12,000 of their activities and objectives to better inform teacher selection and training, noted that as scholars they could only go so far in their recommendations given the American educational context. They conceded, "Since teachers must be trained to produce the sort of citizen that the community expects from the schools, the builder of the teacher-training curricula must first learn the objectives and content of the existing public school curricula and then construct a curricula that will train teachers to reach the objective determined" (pp. 9-10). As a result, "each type of training school must determine its own objectives, and consequently we have not attempted to set up the objectives of public education" (p. 10). At the same time, they bemoaned the "disquieting lack of coordination" in teacher training programs (p. 16).[21] The lack of coordination among and within teacher education programs can be understood as both a cause and a consequence of the absence of two other pieces of an educational infrastructure: a specialized, common vocabulary with which to discuss and improve shared practices, and a shared set of professional norms for practice to guide teachers' thinking and assessment with respect to the quality of their work and that of their students. Missing shared norms, shared practices, and a shared language complicated researchers' efforts to communicate with one

---

[21] This lack of coordination among and within teacher education programs has persisted into the present (Levine, 2006), and this problem "varied treatments" continues to vex researchers of teacher quality who attempt to discover whether or not teacher certification of professional development programs result in teachers of higher "quality" than their peers who did not go through such programs.

another and with practitioners about their research and its value, in addition to complicating communication about problems of practice among teachers.

The Gates MET project and other current efforts to measure teacher and teaching quality are not immune to any of these concerns. For example, while their results are quite useful for thinking about ways of measuring different facets of teaching for different purposes, the fact that other states and districts might have different objectives, observation protocols, or assessments limits their utility. The Gates Foundation (2012) therefore explains that despite their belief in the strength of their findings and implications, "we urge states, districts, and others to apply these same tests to their own data on their own measures of teacher performance and to use the findings to inform continuous improvement of these measures" (p.15). What is more, even if particular instruments do prove to be predictive of student achievement across multiple assessments, some communities may balk at their adoption for the purposes of evaluating or developing instructional quality, since the visions of strong teaching articulated in these tools may or may not be aligned with those of the educators or politicians in local contexts.[22]

…

In sum, research on teacher and teaching quality has been conducted in a decentralized environment that has contributed to the absence of an organizing framework around which a knowledge base for teacher or instructional quality and its improvement might be built. These studies often yield the null or conflicting findings that one might expect from studies carried out across enormously diverse settings that are not sustained by the elements of a shared

---

[22] For example, while many of the protocols used in the MET project place an emphasis on students' intellectual engagement and the ways in which they are positioned to make meaning of instruction in the classroom, others may not approve. In a rather extreme example, the approved 2012 platform of the Republican party in Texas states that with respect to education, "We oppose the teaching of Higher Order Thinking Skills (HOTS) (values clarification), critical thinking skills and similar programs that are simply a relabeling of Outcome-Based Education (OBE) (mastery learning) which focus on behavior modification and have the purpose of challenging the student's fixed beliefs and undermining parental authority" (p. 13).

infrastructure. This infrastructure would include a set of exams or other means of measuring progress towards educational goals; quality curricula or curricular frames to support teachers and students as they work towards them; professional education, vocabulary, and norms to help teachers make sense of their students' needs and so they might discuss their problems with colleagues and adjust shared programs and methods accordingly in professionally responsible ways. Its absence, and the subsequent challenges it presents to practitioners and researchers seeking clarity about teacher and teaching quality, or tools for its assessment and improvement, is inextricably linked with the design of the political system in which it is embedded.

The future

For over a century, the empirical study of teacher, and later teaching or instructional quality, has been pursued by those who hope their findings might be used to improve teaching as a practice and a profession, and ultimately, to enhance the efficiency and/or equity of the educational enterprise. The analysis above, which details the ways that key features of the U.S. educational landscape have influenced these efforts, suggests that more or more sophisticated research on teacher and instructional quality may not be enough to move us closer to those goals. Without a way to organize research efforts, findings, or the ways them might be broadly applied, and without addressing some of the methodological and political problems that such a decentralized and idiosyncratic system presents, future research in this area is likely to remain largely disconnected from other scholarly efforts and from practice.

In this context, one promising area for future research in teacher or instructional improvement might be the careful study of educational organizations that have attempted to develop their own somewhat independent subsystems within or at the margins of the traditional,

83

disjointed system of U.S. education. Examining how these organizations (a) define and measure progress towards goals for students; and (b) define teacher or instructional quality and the strategies by which they attempt to learn about, develop, and support quality across classrooms and schools; could indicate (c) the extent to which these organizations or educational subsystems have developed their own educational infrastructures, or elements of it. Some subsystems that would be interesting sites for study already exist (Bryk, 2009; Cohen, 2011; Lampert & Graziani, 2009; Raudenbush, 2009), including some models of the Comprehensive School Reform (CSR) programs discussed above; select charter school networks; somewhat independent teacher development and support systems like Teach For America, or Advanced Placement programs. Although these organizations may be able to bypass some of the constraints of the traditional educational system, they are also likely to be affected by and share concerns with those working to improve public education from more traditional institutional homes. Understanding where these organizations or programs succeed or struggle has the potential to illuminate a great deal about the possibilities and challenges of instructional improvement and reform in the U.S.

References

Allen, D. (Ed.) (1967). *Micro-teaching: A description.* Stanford: Stanford Teacher Education Program.

Allen, M. (2003). Eight questions on teacher preparation: What does the research say? A summary of the findings. Denver, CO: Education Commission of the States.

Angus, D. L. (2001). Professionalism and the public good: A brief history of teacher certification. J. Mirel, (Ed.). Washington, DC: Fordham Foundation.

Ball, D.L., & Bass, H. (2003). Toward a practice-based theory of mathematical knowledge for teaching. In B. Davis & E. Simmt (Eds.), Proceedings of the 2002 Annual Meeting of the Canadian Mathematics Education Study Group, (pp. 3-14). Edmonton, AB: CMESG/GCEDM.

Ball, D.L. & Cohen, D.K. (1999). Developing practice, developing practitioners: Towards a practice-based theory of professional education. In G. Sykes and L. Darling-Hammond (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 3-32). San Francisco: Jossey Bass.

Ball, D.L. & Lampert, M. (1999). Multiples of evidence, time, and perspective: Revising the study of teaching and learning. In E. Lagemann & L.S. Shulman, (Eds.), *Issues in education research: Problems and possibilities* (pp. 371-398). San Francisco: Jossey Bass.

Barr, A.S. (1940a). Measurement of Teaching Ability. *Review of Educational Research, 10*(3), 182-184.

Barr, A.S. (1940b). Recruitment for Teacher Training and Prediction of Teaching Success. *Review of Educational Research, 10*(3), 185-190.

Barr, A.S. (1945a). Some Introductory Comments. *Journal of Experimental Education, 14*(1), 1-5.

Barr, A.S. (1945b). Impressions, Trends, and Further Research. *Journal of Experimental Education, 14*(2), 200-206.

Barr, A.S. (1949). Measurement and Prediction of Teaching Success. *Review of Educational Research, 19*(3), 185-190.

Barr, A.S. (1952). The Measurement of Teacher Characteristics and Prediction of Teaching Efficiency. *Review of Educational Research, 22*(3), 169-174.

Barr, A.S., Eustic, D.E., & Noe, E.J. (1955). The Measurement and Prediction of Teaching Efficiency. *Review of Educational Research 25*(3), 261-269.

Barr, A.S. & Jones, R.E. (1958). The Measurement and Prediction of Teaching Efficiency. *Review of Educational Research 28*(3), 256-264.

Barr, A.S., Torgerson, T.L., Johnson, C.E., Lyon, V.E., & Walvoord, A.C. (1935). The validity of certain instruments employed in the measurement of teaching. In H.M. Walker, (Ed.) *The Measurement of Teaching Efficiency*. New York: Macmillan Co.

Barr, R., & Dreeben, R. (1983). *How schools work.* Chicago: University of Chicago Press.

Bellack, A.A. & Huebner, D. (1960). Teaching. *Review of Educational Research 30*(3), 246-257.

Berliner, D.C. (1979). Tempus educare. In P.L. Peterson and H.J. Walberg (Eds.), *Research on teaching: Concepts, findings, and implications.* Berkeley, CA: McCutchan.

Bidwell, C. (1965). The Schools as a Formal Organization. In J. March (Ed.), *Handbook of Organizations* (pp. 973-1002). Chicago: Rand McNally & Co.

Bidwell, C., & Karsarda, J. (1980). Conceptualizing and measuring the effects of school and schooling. *American Journal of Education, 89*, 401-431.

Boardman, C.W. (1928). *Professional Tests as Measures of Teaching Efficiency.* New York: Teachers College.

Bodovski, K. & Farkas, G. (2007). Do instructional practices contribute to inequality in achievement? *Journal of Early Childhood Research 5*(3), 301-322.

Borman, G.D., Hewes, G.M., Overmann, L.T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research, 73*, 125–230.

Boyce, A.C. (1915). *The Fourteenth Yearbook of the National Society for the Study of Education: Part II Methods for Measuring Teachers' Efficiency.* Chicago, IL: University of Chicago Press.

Boyd, W.L. (1978). The changing politics of curriculum policy making for American schools. *Review of Educational Research, 48*(4), 577-628.

Braun, H.I. (2005). Using student progress to evaluate teachers: A primer on value-added models. Policy Information Centre, Educational Testing Service. Princeton, NJ: Educational Testing Service.

Brookover, W.B. (1945). The Relation of Social Factors to Teaching Ability. *Journal of Experimental Education, 13*(4), 191-205.

Brophy, J.E. (1982). How teacher influence what is taught and learned in classrooms. *The Elementary School Journal, 83*(1), 1-13.

Brophy, J.E. (1988). Research on teacher effects: Uses and abuses. *Elementary School Journal, 89(*1), 3-21.

Brophy, J.E., & Good, T.L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 318-375). New York: Macmillan.

Brown, G. (1975). *Microteaching: A Programme of Teaching Skills.* London: Methuen.

Bryk, A., Lee, V., & Holland, P. (1993). *Catholic schools and the common good.* Cambridge, MA: Harvard University Press.

Bryk, A., Sebring, P., Allensworth, E., Luppescu, S., and Easton, J.Q. (2010). *Organizing schools for improvement: Lessons from Chicago.* Chicago, IL: University of Chicago Press.

Burnaford, G. (2001). Teachers' work: Methods for researching teaching. In G. Burnaford, J. Fischer, & D. Hobson (Eds.), *Teachers Doing Research: The Power of Action Through Inquiry* (2nd ed., pp. 49-82). Mahwah, NJ: Lawrence Erlbaum Associates.

Carnoy, M. & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. Educational Evaluation and Policy Analysis 24(4), 305-331.

Charters, W.W. & Waples, D. (1929). *The Commonwealth teacher-training study.* Chicago: University of Chicago Press.

Chetty, R., Friedman, J.N., and Rockoff, J.E. (2011). The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood. NBER Working Paper No. 17699.

Clark, C., & Lampert, M. (1986). The study of teacher thinking: Implications for teacher education. *Journal of Teacher Education, 37*(5), 27-31.

Clark, C. & Peterson, P.L. (1986). Teachers' thought processes. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 255-296). New York: Macmillan.

Cochran-Smith, M. (2005).The new teacher education: For better or for worse? *Educational Researcher, 34*(7), 3–17.

Cochran-Smith, M., & Lytle, S. L. (1990). Research on teaching and teacher research: The issues that divide. *Educational Researcher, 19*, 2-11

Cohen, D.K. (1995). What is the system in systemic reform? *Educational Researcher, 24*(9), 11-17, 31.

Cohen, D.K. (2005). Professions of human improvement: Predicaments of teaching. In M. Nisan & O Schremer (Eds.), *Educational Deliberations* (pp. 278-294). Jerusalem: Keter Publishers.

Cohen, D.K. (2011). *Teaching: Practice and its Predicaments*. Cambridge, MA: Harvard University Press.

Cohen, D. K. & Ball, D. L. (1999). Instruction, capacity, and improvement. CPRE Research Report No. RR-043. Philadelphia, PA: University of Pennsylvania, Consortium for Policy Research in Education (CPRE).

Cohen, D.K., Moffitt, S.L., & Goldin, S. (2007). Policy and practice: The dilemma. *American Journal of Education, 113,* 515-548.

Cohen, D., Raudenbush, S., & Ball, D. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis, 25*(2), 1-24.

Cohen, D.K. and Spillane, J. (1992). Policy and practice: The relations between governance and instruction. In G. Grant (Ed.), *Review of Research in Education, 18* (pp. 3-49). Washington, DC: American Educational Research Association.

Coleman, J.S., Campbell, E.Q., Hobson, C.J., McPartland, J., Mood, A.M., Weinfeld, F.D., & York, R.L. (1966). *Equality of educational opportunity.* Washington, DC: U.S. Government Printing Office.

Committee on the Criteria of Teacher Effectiveness. (1952). Report of the Committee on the Criteria of Teacher Effectiveness. *Review of Educational Research, 22*(3), 238-263.

Committee on the Criteria of Teacher Effectiveness. (1953). Second Report of the Committee on Criteria of Teacher Effectiveness. *Journal of Educational Research, 46*(9), pp. 641-658.

Cooley, W.W. & Leinhardt, G. (1980). The Instructional Dimensions Study. *Educational Evaluation and Policy Analysis, 2(*1), 7-25.

Corcoran, S., Jennings, J., & Beveridge, A. (2011). Teacher effectiveness on high and low stakes tests. Unpublished manuscript.

Correnti, R. & Rowan, B. (2007). Opening up the black box: Literacy instruction in Schools Participating in Three Comprehensive School Reform Programs. *American Educational Research Journal, 44*(2), 298-338.

Crabbs, L.M. (1925). Measuring Efficiency in Supervision and Teaching. New York: Teachers College.

Cureton, E.E. (1937). The Accomplishment Quotient Technic [sic]. *Journal of Experimental Education 5*(3), 315-326.

Darling-Hammond, L. (1999). Teacher quality and student achievement: A review of state policy evidence. Seattle: Center for Teaching Policy, University of Washington.

Darling-Hammond, L., & Wise, A. E. (1985). Beyond standardization: State standards and school improvement. *Elementary School Journal, 85*(3), 315–356.

Darling-Hammond, L., & Youngs, P. (2002). Defining "highly qualified teachers": What does "scientifically-based research" actually tell us? *Educational Researcher, 31*(9), 13-25.

DeAngelis, K.J., White, B.R., & Presley, J.B. (2010). The Changing Distribution of Teacher Qualifications Across Schools: A Statewide Perspective Post-NCLB *Education Policy Analysis Archives, 18(28).* Retrieved [June 20, 2012], from http://epaa.asu.edu/ojs/article/view/722

Domas, S.J. & Tiedman, D.V. (1950) Teacher Competence: An Annotated Bibliography. *Journal of Experimental Education*, *19*(2), 110-217.

Doyle, W. (1977). Learning the classroom environment: An ecological analysis. *Journal of Teacher Education, 28*, 51-55.

Duncan, M.J. & Biddle, B.J. (1974). *The Study of Teaching.* New York: Holt, Rinehart, & Winston.

Elmore, R.F. (1996). Getting to scale with successful educational practices. In S.H. Fuhrman & J. O'Day (Eds.), *Rewards and Reform: Creating Educational Incentives that Work* (pp. 294-329). San Francisco: Jossey-Bass.

Erickson, F. (1986). Qualitative methods in research on teaching. The cultures of teaching. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 119-161). New York: Macmillan.

Evertson, C., & Green, J.L. (1986). Observation as inquiry and method. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 162-213). New York: Macmillan.

Feiman-Nemser, S., & Floden, R. (1986). The cultures of teaching. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 505-526). New York: Macmillan.

Freeman, D. & Porter, A. (1989). Do Textbooks Dictate the Content of Mathematics Instruction in Elementary Schools? *American Educational Research Journal, 26*(3), 403-421.

Gage, N.L. (1963). Paradigms for research on teaching. In N.L. Gage (Ed.), *Handbook of research on teaching* (2nd ed., pp. 94-141). Chicago: Rand McNally.

Gage, N.L. (1968). An analytical approach to research on instructional methods. *Journal of Experimental Education, 37*(1), 119-125.

Gage, N.L. (1994). The scientific status of research on teaching. *Educational Theory, 44*(4), 371-383.

Gage, N.L., & Needels, M.C. (1989). Process-product research on teaching: A review of criticisms. *The Elementary School Journal, 89*, 253-300.

Gage, N. L. and Orleans, J.S. (1952). Guiding Principles in the Study of Teacher Effectivness. *Journal of Teacher Education, 3,* 294-298.

Gansle, K., Noell, G., & Burns, J. (2012). Do student achievement outcomes differ across teacher preparation programs? An analysis of teacher education in Louisiana. *Journal of Teacher Education*.

Gates, B. and Gates, M. (2011, October 22). Grading the Teacher. *The Wall Street Journal.* Retrieved June, 2012.

Gates Foundation. (2012). Gathering feedback for teaching. Bill and Melinda Gates Foundation.

Glazerman, S., Mayer, D., & Decker, P. (2006). Alternative routes to teaching: The impacts of Teach for America on student achievement and other outcomes. *Journal of Policy Analysis and Management, 25,* 75–96.

Goe, L. (2007). *The link between teacher quality and student outcomes: A research synthesis.* Washington, DC: National Comprehensive Center for Teacher Quality.

Goe, L. (2007b). Linking teacher quality and student outcomes. In C.A. Dwyer (Ed.), *America's Challenge: Effective Teachers for At-Risk Schools and Students*. National

Comprehensive Center for Teacher Quality. Retrieved June, 2007 from: http://www.tqsource.org/publications/NCCTQBiennialReport.php

Goldhaber, D.D., & Brewer, D.J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis, 22*(2), 129–146.

Goldhaber, D. and M. Hansen. (2010). Assessing the Potential of Using Value-Added Estimates of Teacher Job Performance for Making Tenure Decisions.Working paper 31. Denver, CO: Center for Reinventing Public Education.

Grossman, P. L. (1989). A study in contrast: Sources of pedagogical content knowledge for secondary English. *Journal of Teacher Education, 30*, 24 -31.

Grossman, P., Loeb, S., Cohen, J., Hammerness, K., Wyckoff, J., Boyd, D., and Lankford, H. (2010). Measure for Measure: The relationship between measures of instructional practice in middle school English language arts and teachers' value-added scores. Center for the Analysis of Longitudinal Data in Education Research.

Gutstein, E., Lipman, P., Hernández, P., & de los Reyes, R. (1997). Culturally relevant mathematics teaching in a Mexican American context. *Journal for Research in Mathematics Education, 28*, 709–737.

Hamre, B.K., & Pianta, R.C. (2007). Learning opportunities in preschool and early elementary classrooms. In R.C. Pianta, M.J. Cox, & K.L. Snow (Eds.), *School readiness and the transition to kindergarten in the era of accountability* (pp. 49-83). Baltimore: Paul H. Brookes Publishing Company.

Hanushek, E.A. (2002). Teacher Quality. In L.T. Izumi & W.M. Evers (Eds.), *Teacher Quality* (pp. 1-12). Stanford, California: Hoover Institution Press.

Hanushek, E.A. & Rivkin, S.G. (2006). Teacher Quality. In E.A. Hanushek & F. Welch (Eds.), *Handbook of the Economics of Education, Volume 2* (pp. 1052-1078). Amsterdam: Elsevier North-Holland.

Hellfritzsch, A.G. (1945). Factor Analysis of Teacher Abilities. *Journal of Experimental Education, 14*(2), 166-198.

Hess, F. M., Rotherham, A. J., & Walsh, K. (Eds.). (2004). *A qualified teacher in every classroom: Appraising old answers and new ideas.* Cambridge, MA: Harvard Education Publishing Group.

Hiebert, J., Gallimore, R., and Stigler, J.W. (2002). A Knowledge Base for the Teaching Profession: What Would It Look Like and How Can We Get One? *Educational Researcher 31,* 3-15.

Hill, C.W. (1921). The Efficiency Ratings of Teachers. *The Elementary School Journal, 21*(6), 438-443.

Hughes, M.M. and The Staff of the Provo Merit Study. (1958). Teaching Is Interaction. *The Elementary School Journal, 58*(8), 457-464.

Jackson, K.C. (2011). Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence From Teachers. NBER Working Paper No. 15990

Jayne, C.D. (1945). A Study of the relationship between teaching procedures and educational outcomes. *Journal of Experimental Education 14*(2),101-134.

Jerald, C. (2012). Movin' it and Improvin' it. Using both education strategies to increase teaching effectiveness. Center for American Progress.

Johnson, F., Zhou, L., and Nakamoto, N. (2011). *Revenues and Expenditures for Public Elementary and Secondary Education: School Year 2008–09 (Fiscal Year 2009)* (NCES

2011-329). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved [June 20, 2012] from http://nces.ed.gov/pubsearch.

Jones, R.D. (1946). The Prediction of Teaching Efficiency from Objective Measures. *Journal of Experimental Education, 15*(1), 85-100.

Junker, B.W., Wiesberg, J., Matsumura, L.C., Crosson, A., Wolf, M.K., Levison, A., & Resnick, L. (2006). Overview of the Instructional Quality Assessment. (CSE Technical Report # 671). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Kane, T., Rockoff, J.E. and Staiger, D.O. (2006). What Does Certification Tell Us About Teacher Effectiveness? Evidence from New York City. NBER Working Paper #12155.

Kirst, M.W. and Walker, D.F. (1971). An analysis of curriculum policy-making. *Review of Educational Research, 41*(5), 479-509.

Knight, F. B. (1922). *Qualities Related to Success in Teaching.* New York: Teachers College.

Kupermintz, H. (2003). Teacher effects and teacher effectiveness: A validity investigation of the Tennessee value added assessment system. *Educational Evaluation and Policy Analysis, 25*(3), 287-298.

Ladson-Billings, G. (1995a). But that's just good teaching! The case for culturally relevant pedagogy. *Theory Into Practice, 34*(3), 159–165.

Ladson-Billings, G. (1995b). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal, 32*(3), 465-491.

LaDuke, C.V. (1945). The Measurement of Teaching Ability: Study Number Three. *Journal of Experimental Education, 14*(1), 75-100.

Lagemann, E. C. (2000). *An elusive science: The troubling history of educational research.* Chicago: The University of Chicago Press.

Lampert, M. (1985). How do teachers manage to teach? Perspectives on problems in practice. *Harvard Educational Review, 55,* 178-194.

Lampert, M. (1986). Knowing, doing, and teaching multiplication. *Cognition and Instruction, 3,* 305-342.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal, 27*(1), 29-63.

Lampert, M. (2001). *Teaching problems and the problems of teaching.* New Haven, CT: Yale University Press.

Lampert, M., & Graziani, F. (2009). Instructional activities as a tool for teachers' and teacher educators' learning. *Elementary School Journal,* 109(5), 491–509.

Lee, C.D. (2001). Is October Brown Chinese? A Cultural Modeling Activity System for Underachieving Students. *American Educational Research Journal, 38*(1), 97-141.

Lee, V. E. and Burkam, D. T. (2002). *Inequality at the Starting Gate: Social Background Differences in Achievement as Children Begin School.* Washington D.C.: Economic Policy Institute

Leinhardt, G., & Greeno, J. G. (1986). The cognitive skill of teaching. *Journal of Educational Psychology, 78*(2), 75-95.

Leinhardt, G., Zigmond, N., & Cooley, W. (1981). Reading instruction and its effects. American *Educational Research Journal, 18*, 343–361.

Lieberman, A. (2000). Networks as learning communities: Shaping the future of teacher development. *Journal of Teacher Education, 51*(3), 221-227.

Lipscomb, S., Teh, B., Gill, B., Chiang, H., & Owens, A. (2010). Teacher and principal value-added: Research findings and implementation practices. Cambridge, MA: Mathematica Policy Research.

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

McCaffrey, D.F., T.R. Sass, J.R. Lockwood, and K. Mihaly. "The Inter-Temporal Variability of Teacher Effect Estimates." *Education Finance and Policy,* vol. 4, no. 4, 2009, pp. 572–606.

McCaffrey, D.F., Lockwood, J.R., Koretz, D.M., & Hamilton, L.S. (2003). Evaluating value-added models for teacher accountability, MG-158-EDU. Santa Monica, CA: RAND.

McLaughlin, M. W., & Talbert, J. E. (1993). Contexts that matter for teaching and learning. Stanford, CA: Center for Research on the Context of Secondary School Teaching.

Medley, D.M., & Mitzel, H.E. (1959). Some behavioral correlates of teacher effectiveness. *Journal of Educational Psychology, 50*, 239-246.

Medley, D.M., & Mitzel, H.E. (1963). Measuring Classroom Behavior by Systematic Observation. In N.L. Gage (Ed.), *Handbook of research on teaching* (2nd ed., pp. 247-328). Chicago: Rand McNally.

Meriam, L. (1906). *Normal School Education and Efficiency in Teaching.* Teachers College, New York.

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

Milesi, C., & Gamoran, A. (2006). Effects of class size and instruction on kindergarten achievement. *Education Evaluation and Policy Analysis, 28*(4), 287-313.

Mintrop, Heinrich. (2004). High-Stakes Accountability, State Oversight, and Educational Equity. *Teachers College Record 106*, 2128-2145.

Nocera, J. (2012, May 21). Gates puts the focus on teaching. *New York Times.* RETRIEVED JULY 1, 2012.

Nuthall, G. (2004). Relating classroom teaching to student learning: A critical analysis of why research has failed to bridge the theory/practice gap. *Harvard Educational Review, 74*(3), 273-306.

Nuthall, G., & Alton-Lee, A. (1990). Research on teaching and learning: Thirty years of change. *Elementary School Journal, 90*(5), 547–570.

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects. *Educational Evaluation and Policy Analysis, 26*, 237-257

Odenweller, A.L. (1936). *Predicting the Quality of Teaching: The Predictive Value of Certain Traits for Effectiveness in Teaching.* Teachers College, New York.

Orleans, J.S., Clarke, D., Ostreicher, L., Standlee, L. (1952). Some preliminary thoughts on the criteria of teacher effectiveness. Journal of Educational Research 45, 641-648.

Papay, J. P. (2011). Different Tests, Different Answers. *American Educational Research Journal*, 48(1):163-193.

Peterson, P.L., & Clark, C.M. (1978). Teachers' reports of their cognitive processes during teaching. *American Educational Research Journal, 15*(4), 555-565.

Phillips, K. (2010). What Does "Highly Qualified" Mean for Student Achievement? Evaluating the Relationships between Teacher Quality Indicators and At-Risk Students' Mathematics and Reading Achievement Gains in First Grade. Elementary School Journal, v110 n4 p464-493.

Powell, A. G. (1996). Motivating Students to Learn: An American Dilemma. In S. H.

Fuhrman and J. A. O'Day (Eds.), *Rewards and Reform: Creating Educational Incentives That Work*. San Francisco: Jossey-Bass.

Purkey, S. C., & Smith, M. S. (1983). Effective schools: A review. *Elementary School Journal, 83*(4), 427–452.

Putnam, R.T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher, 29*(1), 4–15.

Raudenbush, S. (2009). The Brown legacy and the O'Connor challenge: Transforming schools in the image of children's potential. *Educational Researcher 38*(3):169-180.

Raudenbush, S.W., Bryk, A.S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*, Second Ed. Sage: Thousand Oaks, CA.

Rice, J.M. (1893). *The Public-School System of the United States.* New York: The Century Co.

Rice, J. (2003). *Teacher Quality: Understanding the Effectiveness of Teacher Attributes.* Washington, DC: Economic Policy Institute.

Rivers, J.C. and Sanders, W.L. (2002). In L.T. Izumi and W.M. Evers (Eds.), *Teacher Quality.* (pp. 13-23). Stanford, California: Hoover Institution Press.

Rivkin, S.G., Hanushek, E.A., & Kain, J.F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417-458.

Rockoff, J.E., Speroni, C. (2010). Subjective and objective evaluations of teacher effectiveness. *American Economic Review* 100:261-266.

Rolfe, J.F. (1945). The Measurement of Teaching Ability: Study Number Two. *Journal of Experimental Education* 14(1), pp.52-74.

Rosenshine, B.V. (1970). Evaluation of Classroom Instruction. *Review of Educational Research*, *40*(2), 279-300.

Rosenshine, B.V. (1983). Teaching functions in instructional programs. *Elementary school Journal, 83,* 335-352.

Rosenshine, B.V., & Stevens, R. (1986). Teaching functions. In M. C. Whittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 376-391). New York: Macmillan.

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom.* New York: Holt, Rinehart, & Winston.

Rostker, L.E. (1945). The Measurement of Teaching Ability: Study Number One. *Journal of Experimental Education, 14*(1), 6-51.

Rothstein, J. (2008). Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement. NBER Working Paper No. W14442.

Rowan, B. (1990). Commitment and control: Alternative strategies for the organizational design of schools. *Review of Research in Education, 16*, 353–389.

Rowan, B. (2002). The ecology of school improvement: Notes on the school improvement industry in the United States. *Journal of Educational Change, 3*, 283-314.

Rowan, B., Camburn, E., and Barnes, C. (2004). Benefiting from comprehensive school reform: A review of research on CSR implementation. In C. Cross (Ed.), *Putting the pieces together: Lessons from comprehensive school reform research* (pp. 1-52). Washington, DC: National Clearinghouse for Comprehensive School Reform.

Rowan, B., Correnti, R., and Miller, R.J. (2002). What large-scale survey research tells us about teacher effects on student achievement: Insights from the Prospects Study of elementary schools. *Teachers College Record, 104*(8), 1525-1567.

Rowan, B., & Miller, R. J. (2007). Organizational strategies for promoting instructional change: Implementation dynamics in schools working with comprehensive school reform

providers. *American Educational Research Journal, 44*(2), 252-297.

Ruediger, W.C. & Strayer, G.D. (1910). The qualities of merit in teachers. *Journal of Educational Psychology, 1*(5) 272-278.

Ryans, D.G. &Wandt, E. (1952). A Factor Analysis of Observed Teacher Behaviors in the Secondary School: A Study of Criterion Data. *Educational and Psychological Measurement, 12,* 574-586.

Sanders, W.L., & Rivers, J. C. (1996). Cumulative and residual effects of teachers on future student academic achievement. Knoxville: University of Tennessee Value-Added Research and Assessment Center.

Sass, T. (2008). The stability of value-added measures of teacher quality and implications for teacher compensation policy. National Center for the Analysis of Longitudinal Data in Education Research Issue Brief #4.

Seidel, T. & Shavelson R.J. (2007). Teaching Effectiveness Research in the Past Decade: The Role of Theory and Research Design in Disentangling Meta-Analysis Results. *Review of Educational Research, 77*(4), 454-499.

Shavelson, R.J., Webb, N.M., & Burstein, L. (1986). The measurement of teaching. In M.C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 55–91). New York: Macmillan.

Shulman, L.S. (1986a). Paradigms and research programs in the study of teaching: A contemporary perspective. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 2-36). New York: Macmillan.

Shulman, L.S. (1986b). Those who understand: Knowledge growth in teaching. Educational Researcher, 15(2), 4-14.

Smith, M.S. & O'Day, J. (1990). Systemic School Reform. In *Politics of Education Yearbook, 1990* (pp. 233-267). New York: Taylor and Francis.

Stallings, J. (1975). Implementation and Child Effects of Teaching Practices in Follow Through Classrooms. *Monographs of the Society for Research in Child Development, 40*(7/8), 1-133.

Stodolsky, S. (1988). *The subject matters: Classroom activity in math and social studies.* Chicago: University of Chicago Press.

Taylor, B. M., Pearson, P. D., Peterson, D. S., & Rodriguez, M. C. (2003). Reading growth in high poverty classrooms: The influence of teacher practices that encourage cognitive engagement in literacy learning. *Elementary School Journal 104*(1), 3-28.

Texas Republican Party. (2012). 2012 Republican Party of Texas Report of Platform Committee and Rules Committee. Retrieved July 1, 2012 from http://convention.texasgop.org/

Thorndike, E.L. (1912). The Measurement of Educational Products. *The School Review, 20*(5), 289-299.

Thorndike, E.L. (1929). The Need of Fundamental Analysis of Methods of Teaching" *The Elementary School Journal*, 30(3), pp. 189-191.

Tomlinson, L.R. (1955). Pioneer Studies in the Evaluation of Teaching. *Educational Research Bulletin, 34*, 63-71.

Torgerson, T.L. (1934). The Measurement and Prediction of Teaching Ability. *Review of Educational Research, 4*(3), 261-266.

Tyack, D. (1974). *The one best system: A history of American urban education.* Cambridge, MA: Harvard University Press.

U.S. Congress. (2001). No Child Left Behind Act. Public Law 107-110. January 8, 2002.

U.S. Department of Education, Digest of Education Statistics. (2011). Table 85: Staff employed in elementary and secondary schools, by functional area: Selected years, 1949-50 through fall 2009.

Walker, H.M. (Ed.) (1935). *The Measurement of Teaching Efficiency*. New York: Macmillan Company.

Waller, W. (1932). *The sociology of teaching*. New York: John Wiley.

Walsh, K. (2001a). Teacher certification reconsidered: Stumbling for quality. Baltimore, MD: Abell Foundation.

Walsh, K. (2001b). Teacher certification reconsidered: A rejoinder. Baltimore, MD: Abell Foundation.

Weick, K. (1976). Educational Organizations as Loosely Coupled Systems. *Administrative Science Quarterly, 21*(1), 1-19.

Weick, K. & McDaniel, R., (1989). How Professional Organizations Work: Implications for School Organization and Management. In T.J. Sergiovanni & J.H. Moore (Eds.), *Schooling for Tomorrow: Directing Reforms to Issues that Count* (pp. 330-355). Boston, MA: Allyn and Bacon.

Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 315-327). New York: Macmillan.

Wilson, S.M., & Floden, R. E. (2003). Creating effective teachers: Concise Answers for Hard Questions. An Addendum to the Report *Teacher preparation research: Current knowledge, gaps, and recommendations*. Education Commission of the States. (ERIC Document Reproduction Service No. ED 476366).

Wilson, S.M., Shulman, L.S. and Richert, A.E. (1987). '150 different ways' or knowing: Representations of knowledge in teaching. In J. Calderhead (Ed.) *Exploring Teachers' Thinking* (pp. 104-124). London: Cassell Education.

Wineburg, S. S. (2001). *Historical thinking and other unnatural acts: Charting the future of teaching the past.* Philadelphia: Temple University Press.

Wittrock, M. C. (1986). Students' thought processes. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 297-314). New York: Macmillan.

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11*, 57–67.

Xu, Z. Ozek, U., and Corritore, M. (2012). Portability of Teacher Effectiveness Across School Settings. CALDER Working Paper No. 77. National Center for Analysis of Longitudinal Data in Education Research and American Institutes for Research.

CHAPTER THREE

INDIVIDUALISTIC AND SYSTEMIC APPROACHES TO TEACHING QUALITY: BUILDING PROFESSIONAL CAPABILITY IN A CHARTER SCHOOL NETWORK

Several decades of research have firmly established that teachers vary widely in their success at helping students to learn, and that this variation is large and consequential for students' academic trajectories (e.g., Chetty, Freidman, & Rockoff, 2011; Rivkin, Hanushek, & Kain, 2005; Sanders & Rivers, 1996). This inconsistency is especially damaging for poor or minority students who are more likely to depend on schools for the resources and support necessary for realizing educational and economic success (Ladson-Billings, 2006). As a result, reformers have offered a vast set of policy proposals and program designs for raising the quality and reliability of students' opportunities to learn throughout the United States. But although such reforms seem acutely necessary for advancing the equity and efficiency of the educational system, our knowledge of how to promote large-scale improvement in teaching quality remains limited.

Traditionally, reformers have attempted to improve the quality of students' learning opportunities throughout the educational system by enhancing the qualifications, knowledge, or skills of the *individual* teachers within it. Efforts to improve teacher recruitment and retention, to enhance traditional teacher education or professional development programs, and to rapidly identify and dismiss the system's least effective teachers are typical examples of this approach. But while the expertise that individual teachers bring to their work is immeasurably important, other scholars and practitioners have focused their attention on more *systemic* approaches to developing both individual and collective capability. "Systemic" can mean many things in education reform, but in this essay I use the term to refer to strategies for teacher and teaching

quality improvement that aim to: (a) identify, design, or organize social, material, and technical resources that may be used to support teaching as a practice and profession; (b) promote coherence among these resources; and (c) capitalize upon the deep interdependence of instruction and the conditions in which it takes place. Together, these approaches to teaching quality improvement would (ideally) help to support teachers and the systems in which they work in extending better and more consistent learning opportunities to their students.

In this article I examine teaching quality development and management strategies within Achievement First (AF), a high-performing charter management organization that aims to provide poor and minority students with a reliably strong education. Building upon the success of a single charter school founded in 1999, AF opened its second school in 2004 but operated a network of 17 schools in Connecticut and New York by 2009-2010, the year of this study. AF's swift growth and central commitment to student achievement focused its leaders' attention on teacher and teaching quality across the network, and their schools' relative autonomy as charters enabled them to experiment with novel ways to organize professional support and supervision.

By investigating how AF worked to define, coordinate, and develop teaching quality to attain ambitious goals for student achievement throughout its rapidly growing network, this study offers insight about both individualistic and systemic strategies for supporting consistency in instructional quality; key aspects of the work required to design, build, and maintain performance-oriented educational systems; new ways of organizing educators' work; how a system might be organized to learn to do this work in a context with few models to inform it; and how the environment in which an educational system is embedded may function to support or constrain teaching and learning. I also describe several of the internal and external challenges that faced AF in this work that may limit the organization's ability maintain and enrich the

quality of teaching and learning within the network as it expands, and to directly support similar

reforms in the broader educational system.

THEORETICAL FOUNDATION

Despite a great deal of consensus that the substantial variation in teachers' effectiveness is

damaging to many of their students and to the flourishing of the nation, scholars and practitioners

fiercely debate the best course for ensuring that students have more consistent access to high-

quality opportunities to learn. To many observers, the central argument appears to be over

whether or not teaching quality improvement is most likely to be realized through policies that

focus on either (a) an overhaul of teacher education and professional development programs, or

(b) changes to teacher recruitment, evaluation, retention, and compensation policies designed to

alter the composition of the teaching force (see, e.g., Jerald, 2012). However, a more important

distinction between strategies for improving teaching may be between individualistic and more

systemic strategies. Below, I explore these ideas in more depth before arguing that high

performing charter school networks are particularly useful settings for exploring some of the

possibilities and limitations associated with individualistic and systemic approaches.

The organization and work of teaching: Implications for reform

The U.S.' public education system is famous for its decentralized and disjointed character (see,

e.g., Boyd, 1978; Elmore, 2004; Meyer & Rowan, 1978; Schmidt, Houang, & Shakrani, 2009;

Weick, 1976), which is due in large part due to the design of the political system in which it is

embedded (Cohen & Spillane, 1992; Cohen, 2011). Authority over education is divided across

several levels and branches of government, and among the 50 states, rather than being unified by

a central governmental agency or professional organization. Within states, essential decisions about education have frequently been delegated to school districts, schools, and individual classrooms. There is little substantive coherence among teacher education programs, even within states; educators are also subject to pressures extended by powerful textbook publishers, testing companies, citizen groups, and other nongovernmental entities that have entered the fray.

Teachers have inconsistent access to this torrent of instructional guidance and to other resources—including their colleagues, given the isolation and weak supervision typical of their work (Lortie, 1975; Weick, 1976)—which are then interpreted and enacted by teachers in distinctive ways. While teachers need to exercise at least a moderate degree of professional discretion if they are to teach in ways that are responsive to individual students and that foster their engagement and understanding of academic content (Lampert & Graziani, 2009; Lipsky, 1983), the loose coupling of the U.S. system of public education magnifies the scope and potential liabilities of these judgments. As a result, teachers' instruction and their students' learning varies wildly within and across schools, even when using common curricular materials, assessments, or standards (Brophy, 1982; Freeman & Porter, 1989; Rowan & Correnti, 2009).

Raudenbush (2009) argues that within this fragmented and largely incoherent system, U.S. teachers' frequent assertions of autonomy constitute a reasonable response to the uncertainty they face in their work and have contributed to powerful notions of teachers' work as "*privatized, idiosyncratic practice*" (p. 172). This paradigm—and the nature of the education system that contributed to its dominance—is useful for understanding why individualistic strategies have dominated in the policy sphere as concerns about variable teacher quality have come to the fore of education reform. If skilled teaching is understood to be an individual trait, and if teachers frequently operate as isolated and relatively autonomous individuals within

schools, then individual teachers seem to be a logical focus for efforts to improve student learning system-wide.

Individualistic strategies for teacher and teaching quality reform

Over the past several decades, a set of compelling studies have made the case that traditional strategies for establishing some minimum standard of teaching effectiveness—e.g., requiring teachers to complete a teacher education program and state certification process, or providing incentives for teachers to complete master's degrees—have not been widely successful. Indeed, these qualifications have proven to be weak or inconsistent predictors of student learning (Clotfelter, Ladd, & Vigdor, 2007; Goldhaber & Brewer, 2000; Kane, Rockoff, & Staiger, 2006; Phillips, 2010; Rowan, Correnti, & Miller, 2002; Wayne & Youngs, 2003).

Similarly disappointing results have followed from studies of large-scale professional development (PD) programs aimed at developing the knowledge and skill of practicing teachers. Much of the PD offered across the country has been criticized for being too brief, superficial, and disconnected from content, not to mention the haphazard ways in which teachers access these programs or their failures to substantively influence instruction and learning (e.g. Ball & Cohen, 1999; Birman et al., 2007; Desimone, Porter, Garet, Yoon, & Birman, 2002; Miles, Odden, Fermanich, & Archibald, 2004). Even PD programs with careful research-based designs and multi-faceted delivery mechanisms have met with little measurable success when attempted across multiple schools and districts (Garet et al., 2008; Garet et al., 2011).[23]

---

[23] For example, two recent experimental studies of PD programs aimed at improving teacher knowledge, practice, and student achievement in second grade reading and seventh grade mathematics found they had no significant impact on student learning. The first study examined the impact of a PD program including a teacher institute, additional seminar days, and in-school coaching on second grade reading teachers. Although participating teachers improved their knowledge of reading instruction, changes to instruction were minor and no changes in student achievement were observed (Garet et al., 2008). The second study focused on a similarly extensive program of PD

Some scholars and practitioners see such findings as calls to reform teacher education, the certification process, and PD, and are making serious attempts at their transformation. But others argue that these lackluster results present clear evidence that the education community has yet to design programs or policies that are successful at enhancing the effectiveness of the existing teaching corps, at least at any sort of broad scale. Until we learn how to *make* good teachers, they reason, we must institute policies designed to recruit people who are more likely to be effective in the classroom, or that swiftly identify and remove those teachers who are not, in addition to any attempts to reform existing teacher quality programs.

In a 2010 report for McKinsey & Company, Auguste, Kihn, and Miller draw upon case studies of the world's highest achieving school systems—Singapore, Finland, and South Korea—to argue that all three share a central strategy for ensuring teaching quality: recruiting all of their teachers from the top-third of their academic systems, and then working to develop and retain them. Pointing to statistics indicating that nearly half of U.S. teachers are drawn from the bottom third of college classes, they argue that designing policies to entice graduates from our own "top-third" to join the teaching ranks must be a fundamental part of any comprehensive strategy for improving teaching quality and for competing with our international counterparts.

Another influential set of arguments in favor of altering the teacher quality distribution through changes to the profession's composition involves identifying and "de-selecting" a school systems' least effective teachers, distinguished by estimates of their value-added to student achievement (VA) (see, e.g., Chetty, Friedman, & Rockoff, 2011; Gordon, Kane, & Staiger,

---

designed for seventh grade mathematics teachers. Slight changes to their classroom instruction were accompanied by no significant change in their mathematics knowledge or students' achievement (Garet et. al, 2011).

2006; Hanushek, 2009).[24] This cadre of researchers and policymakers highlights the profound damage that the lowest-performing teachers can inflict upon their students' learning trajectories. Hanushek (2009) estimates that the students taught by the bottom 5 percent of teachers, as measured by their VA, make approximately two-thirds of the academic progress of their average peers in a given year, and that those paired with the bottom 1 percent of teachers only learn about half as much as the average. He calculates that replacing the bottom 5-10 percent of teachers with "average" teachers could have a powerful effect on students' learning, the U.S.'s standing in international achievement rankings, and the country's gross domestic product. Like the authors of the 2010 McKinsey Report, Hanushek suggests that coupling this "de-selection" with changes to the profession's salary structure may further enhance overall teaching quality by increasing its allure for more competitive applicants and by providing stronger incentives for educators to improve. However, since these policies remain untested within any large school system in the U.S., the ultimate impact of these projections remains unknown.

Despite the enormous differences inherent in these approaches to improving the instructional quality in U.S. schools—by enhancing teachers' capabilities through a dramatic overhaul of teacher education, professional development, and the certification process itself, or by manipulating the composition of the teaching force by transforming recruitment and dismissal policies—both are what Fullan (2010) calls *individualistic strategies,* in that they target the qualifications, knowledge, or skills of the individuals who work as teachers. The basic idea behind both is that improvements at the level of the individual teacher can be aggregated to achieve system-level improvements to teacher quality.

---

[24] Some researchers suggest that these VA estimates be combined with other measures, such as principals' ratings (e.g., Hanushek, 2009) or structured classroom observations and student surveys (e.g., Gates Foundation, 2012) to identify teachers for dismissal or remediation.

However, other scholars argue that the U.S.' fragmented educational system, the interwoven paradigm of privatized idiosyncratic practice, and the nature of skillful teaching suggest just the opposite: that *systemic* change will be necessary for any widespread reform of teaching quality. These observers reason that teachers' uneven effectiveness is not simply due to variation in each teacher's natural ability, motivation, or other personal traits. Instead, widespread variation in practice and outcomes, and the disappointing results of many recent reform efforts, are understood to signal the absence of the conditions that would make skillful performance more likely, or to be symptomatic of the broader context and culture of the U.S.' educational system.[25] From this perspective, an overreliance on individuals as a site for reform is misguided because such pervasive and interdependent problems require a more comprehensive response and because individualistic strategies often ignore important social dimensions involved in the ways that expertise in complex work like teaching is accomplished and shared. From this perspective, widespread changes to teacher and teaching quality will not be successful unless individualistic strategies are combined with more systemic approaches.[26]

---

[25] Indeed, Raudenbush (2009) argues that the failure of many recent educational reforms to reduce racial achievement gaps is due to the fact that policymakers tend to manipulate only the tools available within the paradigm of privatized idiosyncratic practice, including things like conventional resources, accountability policy, or school governance structures. He contends: "The implicit assumption is that such top-down changes will in some unspecified way come to transform interactions in classrooms, leading to improved student learning" (p. 172). Yet without a clear understanding of the kinds of instruction they want to see, an interested observer must "wait to see student outcomes and then infers retrospectively whether something good was happening in the classroom" (p. 172) and then lacks information about what transpired in the classroom that might help others.

[26] One way to think about the importance of individualistic and systemic strategies for ensuring quality practice is to look at an example from a different field. While the training of doctors or nurses is essential for reliably good patient care, so are the systems in which they work. For example, if work shifts are designed in ways that allow or require professionals to work without adequate rest, dangerous patient care errors increase. A more tangible material example involves the clear plastic tubing used for a variety of purposes in hospital patients—including the delivery or removal of medicine, fluids, nutrition, or gases. This tubing is largely interchangeable so nurses can easily make dangerous or even deadly mistakes by attaching one of those materials to an incorrect tube, resulting in problems like the delivery of liquid food or air directly to the bloodstream. Advocates for patients and nurses argue that changes should be made to this tubing so tubes with different functions are incompatible, rendering mistakes of this kind impossible regardless of the "quality" of the nurses on duty (Harris, 2010).

Systemic reform strategies for teacher and teaching quality reform

A systemic approach to teaching improvement may be necessary in part because of the nature of instruction itself. Contemporary definitions of instruction suggest that far from being something teachers do to their students, instruction involves interactions among teachers and students around academic content, within particular contexts, across time (Cohen & Ball, 1999; Cohen, Raudenbush, & Ball, 2003; Elmore, 2004; Lampert, 2001). While individual teachers' expertise matters, so do the contexts in which they work. Cohen, Raudenbush, & Ball (2003) observe, "Teachers and students shape environments by what they notice and how they respond, but environments shape attention and response" (p. 127). From this view, teaching quality cannot be understood as a wholly individual trait, but rather as something generated within complex interactions between the resources to which teachers have access through the systems in which they are embedded, and their use and creation of these resources in practice.

In education, how would a system designed to improve students' access to high-quality instruction be organized and managed? How would this system ensure that teachers can acquire the resources they need to be successful, and that they learn to use these resources or share what they have learned with others? Could the same system also ensure that teachers learn to mediate and exercise judgment about the deployment of these resources given the specific needs of their students and particular contexts of their work in ways that are productive for student learning, and to do so with a great deal of reliability?

Observers have long noted that schools have the ability to influence the social and conventional resources available to teachers and students in their work together (Barr & Dreeben, 1983; Bidwell & Kasarda, 1980; Bryk, Sebring, Allensworth, Luppescu, & Easton, 2010; Ingersoll, 2001; Purkey & Smith, 1983; Waller, 1932). Others have focused on

reorganizing educational systems to better coordinate instructional guidance across schools and classrooms within districts or states, and to hold educators accountable for students' performance (Elmore & Burney, 1997; Smith & O'Day, 1990). However, a variety of theoretical, empirical, and practical work from the last several decades suggests that it may be useful to consider developing and deploying some version of what scholars have called "instructional systems" or "regimes," (Bryk, 2009; Raudenbush, 2009; Wilson, 2008), or the related "infrastructure" for practice (Cohen, 2011), to support widespread improvement to teaching quality. These strategies offer a way of conceptualizing the interdependent resources and practices that might be necessary to move beyond reforms focused on individual teachers, school-by-school reform, or reforms that involve structural changes to the system but do not penetrate to the instructional core.[27] Some instantiations of these ideas also indicate how these systems might shape and balance, but still encourage, the exercise of professional judgment so teachers may adjust practice for their students in educationally responsible ways. Several scholars have described such systems, but they paint somewhat different visions of what building an instructional system or infrastructure for practice might entail and of their potential for improving student learning.

Based on his analysis of eight high-performing Boston-area charter schools, Wilson (2008) argues that they rely too heavily on the valiant efforts of teachers from elite academic

---

[27] Elmore's (1996/2004) essay, "Getting to Scale with Good Educational Practice" takes up the question of why widespread change to fundamental educational processes is so difficult to achieve. He lends support to institutional theorists' argument that within the U.S., large changes to the structures of educational system function to "buffer and assimilate the changing demands of a political and social order that is constantly in flux" (pp. 25-26) while the "core of schooling," or the "basic conceptions of knowledge, of the teacher's and the student's roles in constructing knowledge, and of the role of classroom- and school-level structures in enabling student learning" (p. 11) remain relatively unchanged. He argues that realizing meaningful change to this core beyond a small number of classrooms or schools the U.S. would involve disrupting persistent views of good teaching as an individual trait and harnessing institutional incentives to create conditions that facilitate extensive adult learning. Among others, he proposes several strategies similar to those offered by Bryk (2009), Cohen (2011), and Raudenbush (2009), including the development of external norms to establish a shared vision of high-quality practice and shape professional discussions, assessments, and adjustments of practice. Changing organizational incentives by redesigning educators' work and rewards, such as "encouragement and support, access to special knowledge, time to focus on the requirements of the new task, time to observe others doing it," would facilitate teacher learning (p. 38-39).

backgrounds, in part because of the dysfunctions of the broader system in which they work. He explains that currently, "[These] schools rely on nearly heroic efforts by teachers because they inherit students who have been promoted from grade to grade without mastering essential skills at each grade level," which in turn "requires the teacher to possess unusual analytic skill, agility in shaping the curriculum, personal drive, capacity to engage students, and not least, time" (p. 30). While a small number of charters have used this approach to post achievement results that serve as strong evidence against the inevitability of schools' failures to serve poor and minority students, Wilson argues these dependencies must be addressed if more teachers are to find success and sustainability in their work. He proposes that a "powerful instructional system" would allow committed teachers from a range of backgrounds to achieve results similar to those of these charters. This system might include "placement tests and guides for class formation; a sequential, content-rich curriculum tightly linked to state standards and taught to mastery; frequent electronic assessments; detailed pacing charts, and so on" (p. 36).

Bryk (2009) also advocates for instructional systems, focusing on their promise for organizing individual and collective professional inquiry and learning. He rejects notions of teaching and its improvement that call for either "scripted instruction," in which teachers are asked to faithfully implement externally designed lesson protocols and pacing guides, or opposing ideas that require "each individual teacher, each day, to invent instruction anew" (p. 599). Instead, he contends, "skillful teaching" requires teachers to interpret and quickly respond to "a dynamic interplay of understanding" around students, instructional goals, and pedagogical and other resources (p. 599). Expertise of this kind is developed through "many opportunities to engage in guided practice with others who are more expert," which in turn requires "a *professional learning community organized around a specific instructional system*" [emphasis in

the original] (p. 599). For Bryk, an instructional system specifies what teachers need to know

about their students and their "background knowledge, skills, and interests"; it also includes

"very specific pedagogical practices and social routines and expects automaticity in their use" (p.

600). Those working within such an instructional system would have access to shared concepts

about what constitutes student learning, and to a shared language and evidence base to draw

upon as they discuss and assess progress toward these goals.[28]

Similarly, Raudenbush (2009) maintains that developing instructional systems within

subject areas must be central to educational improvement efforts. Reviewing decades of research

on school effectiveness, he argues that schooling can dramatically reduce racial inequality but

depends upon minority students' consistent access to high-quality instruction. To this end,

Raudenbush suggests schools and school systems organize around "a shared *instructional*

regime" or "a system of assessments and instructional interventions that a community of teachers

share to produce clearly defined aims for student learning." He elaborates:

> Such explicit notions of instruction define the work of teaching, the expertise
> required for classroom success, and the role of incentives and accountability in
> motivating expert instruction. A transformed view of teaching, in turn, redefines
> the role of school leadership in mobilizing and deploying the expertise, materials,
> time, and incentives required for successful enactment of the intended instruction.
> (p. 172)

Raudenbush offers a concrete illustration of these ideas through his profile of the efforts of the

Center for Urban School Improvement (USI) at the University of Chicago. USI's leaders

founded the North Kenwood/ Oakland (NKO) charter school in 1998. They worked with

educators and researchers to develop a school-wide formative assessment system called STEP

(Strategic Teaching and Evaluation of Progress), and used it as a foundation around which they

---

[28] Bryk offers Montessori primary education and Reading Recovery as two examples of "the organizing power of a specific instructional system on the activity of a professional learning community" (p. 600).

built an instructional program in literacy. Students' literacy skills were assessed every 10 weeks, and each level of STEP was linked with carefully selected, required texts and a set of instructional strategies designed to help students to progress. Raudenbush observed,

> [T]he development of STEP and its routine application revealed that instruction could not be left to chance or to the judgment of the singular teacher. Of course, some teachers would do well under such a loose system, but such a system could not ensure quality control for all children. Some children in every class would thrive, but not every child would receive the high level of explicit instruction needed. (p. 177).

He suggests that as teachers develop their practice within such a system, they support it by becoming more skillful enactors, by supporting other teachers' improvement, and eventually by helping to revise the instructional systems themselves to better support students' learning. Other parts of the system (e.g., incentives for school leaders) are organized to support the instructional program, as should be policies and resources in the broader environment. Together, these ideas support a "notion of teaching as *shared, systematic practice*" (p. 172) in contrast to the long dominant idea of privatized, idiosyncratic practice in the U.S.

Cohen's conception of an "infrastructure of practice" overlaps considerably with these visions of instructional systems, both in terms of their content and potential. He argues that the resources that teachers need to skillfully mediate instructional dynamics to facilitate their students' learning are many and varied, and include conventional resources, or "money or the things that money buys" (Cohen, et al., 2003, p. 120). But teachers, engaged as they are in a "practice of human improvement" (Cohen, 2005, 2011)—or efforts to improve and transform other people—are also strongly influenced by the social resources to which they have access. Cohen (2011) argues there are three kinds of interdependent social resources essential for widespread success in human improvement practices: an infrastructure of practice; mutual commitment to improvement; and "the work that practitioners and clients create together" as

they meet (pp. 54-55). In education, Cohen identifies three domains in which teachers and learners meet, and through which teachers might work to connect teaching with learning using social resources created and used across them, including "knowledge and how it is extended, the organization of instructional discourse, and acquaintance with students' knowledge" (p. 40). Mutual commitment is critical to these efforts because of the reciprocal dependence that characterizes the practitioner-client relationship in practices of human improvement. After all, Cohen observes, "Teachers cannot 'learn' students. Students depend on teachers for help, yet teachers' success depends on students' learning" (p. 50). If students and teachers are to engage in the difficult work of teaching and learning together, their mutual commitment is crucial. But such commitment is of limited value without an infrastructure of practice.

Cohen remarks that many elements of an instructional system are more typical in other lines of work, and in fact, make successful practice possible.[29] This does not diminish the importance of individual practitioners; he writes, "Occupations cannot practice—only individuals or groups of individuals can do that," which requires careful attention to individuals' knowledge and skill (p 49). However, he argues, "Skilled occupations cannot thrive without the extensive technical affordances that enable work" (p. 56). Together, these affordances constitute an "infrastructure of practice" (p. 54) and include:

> […] Socially accepted views of the proper terrain of work; the sorts of problems that practitioners can claim to solve and the results that should be expected; the specialized knowledge and skills required to solve those problems; the specialized terminology needed to discuss, plan, perform, and assess work; the education that prepares people to work and to improve; norms and standards that inform judgments about the quality of work; the organized intelligence that influences the invention of new tools and technologies; and procedures to deal with unacceptable work. (p. 56)

---

[29] This kind of infrastructure is also far more common in education in other nations (Cohen, 2011; Fullan, 2010; Maroy, 2008; McKinsey & Company, 2007; Schmidt, Houang, & Shakrani, 2009).

In education, an infrastructure if practice would include elements similar to those identified by Wilson, Bryk, and Raudenbush: exams; curricula or curricular frameworks; teacher education; a specialized vocabulary for discussing teachers' practice and student work; and professional standards or norms. If consistent, well-designed, and thoughtfully deployed, an infrastructure might support individual and collective learning and practice, codifying some aspects of educators' developing knowledge and providing tools for instructional development and management. In this essay, I adopt Cohen's phrase "infrastructure of practice" over "instructional system" because this idea encompasses many of the social and organizational arrangements that surround instructional systems as well as the systems themselves.

In the U.S., a unified infrastructure or instructional system consisting of such instruments and supporting an idea of teaching as a shared, systematic practice has been conspicuously lacking though there have been some notable exceptions. Some schools and school districts have worked to build some elements of an infrastructure or instructional system (Bryk et al., 2010; Elmore, 2004; Elmore & Burney, 1997). One set of recent reform efforts, comprehensive school reform (CSR) programs, generated a small number of educational subsystems that ultimately developed versions of instructional systems operating across as many as 1500 schools nationally (Cohen, Gates, Glazer, Goldin, Peurach, forthcoming). A few CSRs had a notable influence on patterns of school management, instruction and student achievement (Borman, Hewes, Overman, & Brown, 2003; Correnti & Rowan, 2007). However, many CSRs ran into problems related to their status as school-level interventions within existing state and district systems and to the lack of infrastructure in the broader educational context. District adoptions of different curricula thwarted CSR implementation, as did the hiring of principals and teachers not knowledgeable about or committed to their designs. CSRs also struggled with the fiscal and human resource

109

demands associated with inventing, and teaching educators to use, elements of an infrastructure in an environment in which these things were unfamiliar and their school sites numerous and geographically dispersed (Berends, Bodilly, & Kirby, 2002; Cohen et al., forthcoming).

Charter schools, charter management organizations, and systems building

Charter schools—public schools of choice established through performance contracts made with an authorizing group sanctioned by state law—occupy a prominent place in the national education reform conversation.[30] The basic logic of charter schools is straightforward: in exchange for greater accountability to an authorizing agency, charter school leaders are afforded far-reaching autonomy in governance, budgets, and staffing so they may meet local needs in innovative and efficient ways. As a result, charters are expected to create high-quality public school options for families and educators who would otherwise have too few, and incentives for surrounding districts to improve by creating competition for staff, students, and funding.

Two decades after the first charter school was established in Minnesota, progress toward these goals has been mixed. Charter schools have expanded throughout the U.S. with legislation in 41 states, Washington D.C. and Puerto Rico permitting their establishment. Still, their proliferation has not been as rapid as their advocates had hoped. In the 2010-2011 academic year, there were only about 5,275 charter schools, representing 5.4 percent of public schools nationally (NAPCS, 2012). However, there are 96 school districts across the U.S. in which charters enroll at least 10 percent of the student body and some enroll considerably more. In New

---

[30] For example, President Obama has made charter school expansion a centerpiece of his administration's school reform agenda and a major funding priority. His administration successfully pressured states to create legislation supportive of charters and their expansion to compete for the 4.35 billion dollar Race to the Top Fund available for K-12 reform and other funds earmarked for charter expansion and turnaround charter schools (Obama, 2009; Duncan 2009a, 2009b).

Orleans, approximately 70 percent of students enroll in charters. In D.C., Detroit, Kansas City, and Flint, between 32-40 percent of students attend them (NAPCS, 2011).

Even in these relatively small numbers, charter school performance has been erratic. Researchers have documented a great deal of variability in their influence on student learning (Gleason, Clark, Tuttle, Dwoyer, & Silverberg, 2010; CREDO, 2009).[31] Charter management organizations (CMOs), which are umbrella organizations that operate a network of charters, have also been established with hopes of sharing resources and ideas among multiple schools to support greater efficiency, quality, and sustainability.[32] However, these networks and the schools within them also vary greatly in their effectiveness relative to the districts their students would have otherwise attended (CRPE, 2011; Tuttle, Teh, Nichols-Barrer, Gill, & Gleason, 2010).

Nonetheless, a small number of high-performing charters have provided a tantalizing glimpse of their potential by supporting their predominantly low-income and minority students in achieving at high levels with remarkable consistency. High-performing CMOs that aim to provide a reliably strong education to traditionally underserved students across a system of schools may be particularly interesting organizations in which to examine systemic approaches to teaching improvement. Because these CMOs aspire to support their students in achieving at high levels across their network, the development and enactment of a system that can coordinate and support work within and across classrooms and schools might be particularly instrumental.

---

[31] Comparing the effectiveness of charter and traditional public schools is not straightforward (Betts, Hill, et al., 2006; Hoxby, 2009; Hoxby & Murarka, 2008; Tuttle, Clark, & Gleason, 2011). For additional reviews of research on charters' effectiveness, see Angrist, et al., 2011; Betts &Tang, 2008, 2011; Carnoy, Jacobsen, Mishel, & Rothstein, 2005; Miron & Nelson, 2004; and NAPCS, 2009. For recent analyses of charters' effectiveness in major urban areas, see Abdulkadiroglu, et al., 2009; Buckley & Schneider, 2007; Hoxby, Murarka, & Kang, 2009; and Teh, McCullough, & Gill (2010).

[32] The NewSchools Venture Fund (2006) has provided important intellectual and financial support to the CMOs attempting to embody these ideas. Members of the group define CMOs as "nonprofit networks of schools that serve a specific geographic area or set of markets" which are "uniquely positioned to maximize quality and sustainability, while also leading to scale within a targeted geographic area" (p. 2). They contrast CMOs with "affiliated" charters that are more loosely related and thus more likely to vary in quality (p. 3), and with for-profit educational management organizations (EMOs) that "may also face pressure from investors or stakeholders to prioritize financial returns over student outcomes" (p. 4).

In addition, charter networks may have greater potential for leveraging their schools' autonomy as charters by creating stable, alternative environments to shelter this work—the absence of which have threatened similar efforts in traditional public schools (Berends, Bodilly, & Kirby, 2002; NewSchools Venture Fund [NSVF], 2006; Raudenbush, 2009).[33]

CMOs are not immune to the types of environmental pressures that have plagued CSRs and other large-scale reform initiatives. They are dependent upon enabling state legislation; meeting the conditions of the charters that authorize them, and especially their accountability provisions; and maintaining strong relationships with parents and teachers (Hill, Lake, & Celio, 2002). Charter leaders must also manage volatile political environments and pressures from private foundations and other funders if they are to thrive (CRPE, 2007). However, as their advocates hoped, their relative autonomy from state and district mandates may offer their leaders important protections from the instability and incoherence of the broader system even if this flexibility has not *ensured* that charters will offer a better quality education (Merseth, 2009; Wilson, 2008). When coupled with their ability to build new systems rather than reforming existing ones, this independence may provide CMOs with opportunities to organize to support teaching quality in novel ways. CMOs may also foster a higher degree of mutual commitment among educators and students than in many traditional educational systems because they or their families have chosen to lead, teach, or attend the schools rather than being assigned to them, and

---

[33] Raudenbush (2009) pointed out that USI's initial goal was to engage a small number of schools in Chicago's South Side in sustained efforts to improve their literacy instruction. However, "the work was inspiring and frustrating: Inspiring because of what it revealed about children's intellectual energy and potential for dramatic growth, frustrating because systemic norms and bureaucratic rules seemed constantly to get in the way of the ambitious instruction the children needed" (p. 176). Establishing an elementary charter school provided USI with new opportunities: "Because charter schools have relaxed rules, USI had pretty much free rein in designing and running [the school], although within quite limited resources." In particular, "The school was free to shape teacher recruitment, curriculum design, and—particularly important—instructional time, to pursue ambitious intellectual goals for the student body, which was nearly all African American and about 75% low income" (p. 176).

they may leave if they discover they are a poor "fit." This mutual commitment provides CMOs with another powerful social resource they may marshal in pursuit of their goals.

The distinctive resources afforded charter school networks, and the limited number of charters and CMOs that significantly outperform traditional public schools, might suggest they hold few useful lessons for the improvement of the broader U.S. school system. However, I propose that *because* of their somewhat unique circumstances and goals, successful CMOs may provide a useful laboratory for us to examine what it might take to organize for consistency in teaching quality across multiple schools under relatively (theoretically) advantageous conditions, and that what we learn in this laboratory can provide critical insight into what might be involved in creating conditions to support greater reliability in teaching quality in the wider environment.

RESEARCH QUESTIONS

This study examined the ways that CMOs worked to define, and organized to attempt to provide, reliably strong instruction for their predominantly low-income and minority students within and across schools. Three sets of questions guided my inquiry:

(1) What are the key instruments—or the social, curricular, educational, political, or material tools designed in influence instruction—that high-performing CMOs use to define, develop, and manage instructional quality within and across schools? How are they selected, designed, and used, individually or together?

(2) How do CMOs organize to manage their relationships with the environment? How does the environment appear to interact with the organization's design, culture, and operation to influence the instruments and how they are used?

(3) What role does organizational learning appear to play in developing, sustaining, and institutionalizing the network's efforts in these areas?

By considering these questions, I hoped to better understand the sorts of resources that CMOs with a focus on strong student achievement across their network bring to bear on their efforts

with respect to teaching; whether or not these resources reflect individualist or systemic strategies for teaching improvement; the extent to which these resources appear to cohere as an "infrastructure of practice," and to what effect. I also wanted to understand what the *work* of organizing to support teaching quality in these CMOs looked like, and examining the key resources that they were designing and deploying seemed to provide a critical window into this work. I was interested to learn about the ways that CMOs cope with their environments because of prior research detailing the powerful if often contradictory influence of the operating environment on educational systems and efforts to improve them, and because of the notion that charter schools might offer important protections for the people and work occurring within them.

Finally, although I focus more on the response to the third question in a separate essay (Rosenberg, 2012), attending to organizational learning in some part seemed important because CMOs and the people within them are likely to have much to learn if they are to undertake these kinds of work. Many of the elements of an infrastructure of practice are rare or disjointed in the broader U.S. context so CMOs would need to develop many of them as they go with staff members who are unlikely to have had experience with coherent educational systems or notions of teaching as a shared practice. The extent to which network employees are able to learn to do the work required of them—to design elements of an infrastructure of practice, to organize to enact the infrastructure together, to socialize new members, to manage the network's material needs, to negotiate environmental pressures and incentives, and to sustain and institutionalize this learning across the organization—will likely have much to do with the organization's success. And whether or not they systematically engage in productive learning will likely depend on whether or not the network is designed to support it.

DATA AND METHOD

To explore these questions, I designed a study with two main parts. First, I selected five CMOs

in three geographic regions with evidence of solid student learning, a focus on instruction, and

ambitious growth plans during the 2009-2010 academic year for the study. All five CMOs

contacted agreed to participate and to be identified in the study. They include: Achievement First

and Uncommon Schools in the northeast (New York, New Jersey, and Connecticut); The Noble

Network in Chicago, Illinois; and Aspire Public Schools and Rocketship Education in California.

Semi-structured, 60-90 minute interviews with 2-4 leaders from each of the CMOs provided

evidence about the similarities and differences among networks with respect to the research

questions. These interviews took place from February through June of 2010.

Second, I completed a more intensive study of one of these networks, Achievement First.

It is from this case study that the primary data for this article are drawn.[34] AF was selected for

intensive study because leaders from peer organizations, educational philanthropies, and other

partner organizations frequently identified both AF and Uncommon Schools as leaders in the

field in terms of their success with student achievement and the development of professional

growth and management systems. Concerns about the ways in which Uncommon School's more

complex organizational structure might compromise my ability to gain a full picture of the

CMO's instructional support mechanisms (Uncommon is a network of networks) led me to select

AF for in-depth study. Fortunately, AF's leadership team approved the research request,

---

[34] Because the study took place in the spring of 2010 and presents a depiction of AF in the 2009-2010 school year, I wrote the case study using the past tense. I chose to do so because AF is vibrant, rapidly expanding and evolving network, and they have continued to develop and change in important ways since the time of the study. Therefore, although much of the material in the case study reflects current practice, some of it does not and therefore it seemed more accurate to describe everything in the past tense. However, I found that using the past tense, while technically more accurate, made AF and the work in which its employees are engaged sound far less dynamic than it is. For example, to write that AF "organized for improvement" rather than writing that it "organizes for improvement" makes it sound like this was something it did once and no longer does, when really this work is an ongoing accomplishment.

suggesting that the public reports resulting from my work would support AF's mission of sharing what its faculty and staff have learned from their experiences to stimulate widespread reform in other charter and traditional school systems.

The AF study is an embedded, single-case study (Yin, 2009) that includes data collected at the network, school, and classroom levels of the organization—documents and archives; audio-recorded, semi-structured interviews with network leaders, school leaders, and teachers; and handwritten field notes from observations of professional development and classroom settings—to explore the areas of inquiry of the more extensive CMO study in greater depth. I selected this design because I want to understand more about the interplay between what happens at the network, school, and classroom levels; the formal and informal mechanisms that connect them; how what happens within the organization is linked to its environment; and how learning about this work occurs in and by the organization.

More specifically, at the network level of the organization, I conducted fifteen 45-90 minute interviews with AF network leaders. These interviews were focused on the leaders' accounts of their professional history and current position and responsibilities; major features of the definition and management of teaching quality in the network; AF's operating environment; and aspects of what AF calls "knowledge capture and sharing" and collaboration within the organization. I also observed one of two yearly, network-wide professional development days, which was frequently cited within the network as a strong example or performance of the network's commitment to maintaining connectedness, to teachers' professional growth, and to the enthusiasm and urgency that characterize the organization's overall culture.

At the school level, I restricted my investigation to AF's seven elementary schools to limit the resources I requested from the organization, to focus my inquiry enough that I felt

confident I could assess the alignment of network staff and school leader perceptions, and because of my background as an elementary school teacher. School leader interviews focused on topics similar to those in the network interviews, though I also asked them to briefly describe their school context, a typical day in their job, and the ways their school days or weeks were organized for students and teachers. I interviewed either the principal, academic dean(s), or both from each school for a total of thirteen 45-90 minute elementary school leader interviews.

Additionally, I selected two mature (i.e. K-4, within the AF model) elementary schools—one in New York, and one in Connecticut—for further study. In both schools, I conducted more in-depth interviews with school leaders, observed their fourth "data day," one of five days of data analysis and planning following the return of interim assessment data. I also spent a week in each school shadowing school leaders, attending teacher and leader work sessions, and observing in classrooms. I conducted 30-60 minute interviews with thirteen teachers in these two schools whose instruction and work with colleagues I observed. I asked them about their professional backgrounds, their experiences with AF, and the content of my observations. While my observations in these two schools were not extensive enough to allow me to make statements about the inner-workings of all of the elementary schools in the network, and the interviews with teachers in each school did not afford insight into the views of all elementary school teachers, these sources of data did provide very useful illustrations of the ways in which the ideas that surfaced in interviews with network and school leaders might play out in practice and for teachers and students in elementary schools in the two states in which AF operates.

Analysis of all sources of data was ongoing and comparative, allowing my developing hypotheses about how AF and the other networks define and organize to support instructional quality to inform subsequent data collection (Miles & Huberman, 1984). Descriptive and analytic

memos synthesizing information and key insights from documents, artifacts, archives, and field notes collected from direct observations helped to organize the data from these sources while permitting me to document preliminary findings in response to the study's research questions. The interviews were transcribed and coded using Atlas.ti. Although sensitizing concepts from prior research and the data collection process shaped my initial coding scheme of descriptive codes (e.g., interviewees' accounts of their role and responsibilities; their definition of a "good teacher" within AF; their references to life-work balance and sustainability; or their discussion of particular professional development initiatives, etc.), I continued to refine the codes using categories grounded in the data throughout the process of analysis (Strauss & Corbin, 1998). Multiple sources of data support data triangulation to guard against threats to construct validity (Maxwell, 2005; Yin, 2009), and member checks conducted with AF staff members have helped to minimize the bias associated with my outsider status (Lincoln & Guba, 1985).

During the informed consent process associated with the interviews, individuals working for AF's network office were given the option of being identified by name and/or position in reports resulting from the study. When they agreed to be identified, I have used their names, titles, or both here. Those network personnel who preferred not to be identified, school leaders, and teachers are identified by their generic role (e.g., network leader), and an interview number. In a small number of cases, when I believed that a quote might identify an individual and allow their other quotes to be identified as well, I redacted the interview number note this.

ACHIEVEMENT FIRST BACKGROUND

The Amistad Academy, a public charter school in New Haven, Connecticut was founded in 1999 by a small group of reformers and educators determined to demonstrate that when provided with

an excellent education urban students are capable of achieving as highly as their more affluent

suburban peers. Four years later, having led their students to outperform district and state

averages (often dramatically), Amistad's leaders established a non-profit charter school

management organization they named Achievement First (AF) to support the founding of

additional schools that would build upon what they were learning at their flagship school.

AF opened its second school in 2004. By the year of this study, the 2009-2010 academic year,

the network included 17 elementary, middle, and high schools serving 4,500 students in

Bridgeport, Hartford, and New Haven, Connecticut and Brooklyn, New York. These schools

appear to have met or exceeded many of Amistad's early successes.[35] Their students—more than

98 percent of whom were African American or Latino, and 74 percent of whom qualified for a

free or reduced-price lunch—outperformed district and state achievement averages with

remarkable consistency, and students graduating from AF high schools have enrolled in college

at impressive rates. AF plans to maintain the rapid pace of their expansion; the organization's

target growth trajectory would establish a network of 30 schools serving 12,000 students in the

next decade.[36]

---

[35] Though AF students do perform well on achievement tests across their classrooms and schools, as was noted in the literature review, isolating the causal impact of AF' s program on their students' learning can be challenging and as yet no studies that have focused exclusively on the impact of AF's schools have been publicly released. Hoxby, Murarka, & Kane's (2009) study taking advantage of the experimental conditions inherent in oversubscribed charters' lottery systems estimated that New York City charters, on average, had a positive and meaningful effect on both mathematics and reading achievement. They also found positive associations between many features of the educational programs of these charters that characterize AF (extended school day, etc.). Teh, McCullough, and Gill (2010) provide some additional evidence to this effect. Using a variation of a difference-in-differences matching strategy, the researchers identify students with similar characteristics and achievement before beginning middle school. The difference between the academic growth of this matched comparison group before and after beginning middle school and the growth of those who enroll in AF/Uncommon is compared. They found that during the academic years spanning 2005-2008, the five middle schools affiliated with both AF and Uncommon Schools in New York City in the study had a positive influence on student learning in math after two and three years of enrollment, and in reading after three years. These effects translate into approximately 0.9 and 0.7 years of additional achievement in math and reading, respectively.

[36] During the 2011-12 school year, AF operated 20 schools employing 585 teachers and serving approximately 6,200 students in grades K-12 (Doyle & Han, 2012). They plan to expand to a new state, Rhode Island, in 2013.

AF's leaders reported that by expanding the number and geographic reach of their schools, they aspired to establish that achievement gaps by race and class can be closed at the scale of an urban school district, and to do so in ways that are replicable for other charters and traditional districts to inspire broader reform. Meanwhile, they aimed to increase traditionally disadvantaged students' access to schools designed to prepare them "with the academic and character skills they need to graduate from top colleges, to succeed in a competitive world and to serve as the next generation of leaders for our communities" (AF, 2011). AF's leaders have attempted to leverage their schools' relative autonomy as charters to pursue these two fundamental commitments: to provide poor and minority students with the necessary resources to succeed on state assessments and beyond, and to do so at the scale of an urban district.

To allow students to attend an AF school continuously throughout their K-12 education, AF has developed a model for school growth that involves clusters of 5 schools, with 2 elementary schools (K-4) sending students to two middle schools (5-8 grades) and then to a single high school (9-12). Elementary schools generally begin with only students in Kindergarten and first grades, and then grow by a grade a year; middle schools are started with a single fifth grade class that then forms the ninth grade class of a new high school upon their graduation. The AF management organization, originally named the AF Central but recently renamed AF Network Support (AFNS), is an umbrella organization for these schools but it does not manage them directly.[37] Instead, several schools serving different grades within an AF cluster are usually opened as different campuses under a single charter that share a governing board of trustees. These schools apply for charters renewable every five years in Connecticut and New York using

---

[37] Throughout this essay, I use AFNS to refer to the umbrella charter management organization (exclusive of the schools), and use AF to refer to the organization as a whole (including schools).

AF's name and model, and enter into a formal charter management agreement (CMA) with the non-profit AFNS, which has its own board of trustees.

The CMA carefully outlines the legal, financial, educational, and normative aspects of the relationship between AFNS and the network's schools. For example, in exchange for a service fee, AFNS is explicitly charged with providing critical services related to managing the school, including curriculum development, professional development, recruitment of school leaders and teachers, school inspection and evaluation, managing issues of facilities, fundraising, and marketing. Meanwhile, AF schools are responsible for successful day-to-day school operations and share responsibility with AFNS for things like the student recruitment and admissions process. Because either AFNS or the schools are able to terminate the CMA if they decide the other group is not contributing adequately to the partnership, the CMA helps to establish a relationship of mutual accountability between them.[38]

During this period of rapid growth, AF staff members have depended upon a mix of revenue sources to support their work. Schools were funded almost exclusively by per pupil funds allocated by the state and additional public sources of revenue (e.g. Title I funds), and they paid AFNS 8-10 percent of the per pupil monies for their services. Because AF schools begin with only one or two grades of students, as schools are founded and require intensive support from AFNS they do not pay as much to AFNS as when they are fully enrolled. In part because of this, and because of the start-up costs associated with establishing new schools (furniture,

---

[38] These formal agreements reflect AF's vision of the potential of a "network approach." They write that organizing as a network is meant to "support and enhance our schools in our shared mission to close the achievement gap" through the pursuit of five objectives by AFNS: (1) *Freeing schools to focus on achievement*, teaching, and learning, by taking on responsibility for some functions that are easily centralized, such as "teacher recruitment, fundraising, budgeting and fiscal operations, data management, information technology, and facilities operations" and organizing schools so they have operational support as well; (2) Assisting with key *talent development* functions for similar reasons (3) *Knowledge capture and sharing* to share and systematize significant innovation throughout the network; (4) *School support and quality control* via coaching relationships, professional development, networking, and the ability to intervene if things go awry; and (5) *Efficiencies that enable the network to be sustainable*, through economies of scale and the specialization of services (AF, n.d.).

textbooks, etc.) and building the tools AFNS uses to support its schools (e.g., an online assessment platform), AF has also drawn heavily on philanthropic support. The fact that Connecticut charters are afforded approximately 25 percent less per pupil than their host districts added financial challenges for AF, as did the fact that in Bridgeport and New Haven, they incurred facility and/or occupancy costs (finding a building, lease/mortgage fees, and operating and maintenance costs) that were not taken out of non-charter schools' operating budgets. Still, AF's leaders aspired to operate with the same financial resources as their host districts but did not anticipate doing so until more of its schools are established and fully subscribed (AF, 2009b).

Like many other high-performing charter, AF engages in a version of the "No Excuses" approach to schooling common to many high-performing charter schools and networks. Wilson (2008), citing Carter (2001) and Tough (2006), writes that the No Excuses style can be broadly described as follows: "Highly educated, driven, and generally young teachers lead their students in a rigorous academic program, tightly aligned with state standards, that aims to set every child on the path to college" (p. 7). The approach gets its name because "founders and staff steadfastly reject explanations from any quarter for low achievement, whether a district apologist's appeals to demographic destiny or a child's excuse for failing to complete an assignment" (p. 7). Other features of the model include the schools' small size, strict disciplinary system, extended school day and year, and highly selective hiring process which aims to yield strong, empowered school leaders and a smart, driven faculty that is generally non-unionized.

This broad-stroke description roughly fits with AF's core values and educational model. Of AF's nine core values, "no excuses" is the first. Other core values reflect AF's attention to the importance of both individuals and their actions (e.g. people matter, mightily; and everything with integrity) and to the potential of their collective work (e.g. team and family; and many

minds, one mission), while capturing some of the urgency and relentless pursuit of achievement that often characterizes No Excuses schools (e.g. excellence is a habit; sweat the small stuff; first things first—which refers to the needs of students—as does, whatever it takes). AF schools are expected to share six "core program elements" that align with a No Excuses approach and reflect these broader organizational values. In brief, they include (1) an unwavering focus on student achievement; (2) a focus on teaching and school leadership talent; (3) more time on task (the AF school day is nearly two hours longer than the school days in their host districts, and students attend a 15-day Summer Academy each summer); (4) rigorous curriculum; (5) strategic use of data and targeted interventions; and (6) a strong school culture. This school culture is characterized by attention to core character values (or "REACH" values—respect, enthusiasm, achievement, citizenship, and hard work); "sweating the small stuff"; a focus on college; small schools in which "teachers know and care"; an orientation towards "parents as partners"; a focus on attendance; uniforms; and the "joy factor" which is intended to permeate the rigorous and structured environment AF aims to create (AF 2009b, pp. 16-17). The network has organized and elaborated the connections between many of these elements of their model and values through their map of their theory of change (see Appendix A) and articulation of a set of strategies and metrics by which to gauge their progress and success in meeting their goals (see Appendix B for AF Strategy Map and Balanced Scorecard Metrics).

However, this overview of the network and its educational approach only begins to allude to the ways that AF's leaders were working towards ensuring that their teachers were able to reliably support their students in meeting the ambitious goals the network holds for them. Indeed, because teachers are the most important school-based influence on student learning yet typically vary widely in their effectiveness, teaching quality was one of the rapidly growing network's

most central preoccupations and their efforts in this area have structured much of the rest of the organization's work.

FINDINGS

AF's leaders have come to believe that the strength of the individual teachers within the network, the quality of the tools and practices these teachers have at their disposal, and their opportunities to (learn to) use them are all essential for supporting high quality teaching and learning as the organization grows. Therefore, as part of their quest to support consistently strong teaching and learning across their classrooms and schools, AF's leaders have carefully invested in both individualistic teacher quality strategies, such as recruitment, retention, and dismissal policies and professional development aimed at developing each teachers' knowledge and skills, as well as more systemic efforts, including their work to codify and institutionalize a set of systems to establish a shared vision of instructional quality; common network and school practices to support at least minimal competence in this instruction; and roles, relationships, and processes to foster continuous improvement in individual and collective work. AF has borrowed ideas and resources from peer organizations, educational systems in other nations, the business sector, and major funders as they have undertaken this work. Nonetheless, AF's leaders have put the pieces together, and invented new parts and connections among them, in innovative ways. As they were used together within AF in the spring of 2010, these individualistic and systemic approaches offered a strong example of an infrastructure of practice—if still emergent.

One could make a convincing argument that *all* of AF's work contributed in some fundamental way to successful classroom practice and outcomes by helping to build systems to support strong teaching and learning, by improving upon them, or by securing or protecting

essential resources for their benefit. However, in this essay I focus on the three most essential categories of AF's work as it related to instructional quality: (1) recruitment and selection, (2) instructional tools, and (3) instructional leadership and organizational design. Given the dearth of examples of what an infrastructure of practice in the U.S. might look like, I provide a broad descriptive overview of the AF leadership's account of the design of these three areas of AF's labors and the ways in which these resources and their productive use intersected with AF's dominant culture. "Organizational culture" is a widely used but controversial concept with scholars debating the ways it might be understood, defined and studied. Following Sørenson (2002), I borrow O'Reilly and Chatman's (1996) definition of organizational culture as "a system of shared values (that define what is important) and norms that define appropriate attitudes and behaviors for organizational members (how to feel and behave)" (p. 160).[39] Organizational culture "encompasses what is valued, beliefs about how things work, and behavioral norms for how work is carried out" (Vogus, Sutcliffe, & Weick, 2010, p. 60). While AF's culture is not a part of its infrastructure of practice per se, they are mutually constitutive and therefore inextricable, particularly when attempting to understand the ways the infrastructure was used. AF's infrastructure of practice and culture offered affordances for teaching quality development and management rarely found in traditional school systems. However, at the end of the essay I also identify several features of AF's approach and operating environment that may complicate the organization's ability realize its leaders primary goals of supporting poor and minority student success and the sustainability and replication of their efforts.

---

[39] One key debate is the extent to which an organizational culture can be understood to be singular. For example, Martin (2002) identifies three theoretical views on organizational culture in the research literature: an integration perspective that emphasizes a consensus view of culture; a differentiation perspective, which rejects a monolithic view of culture and focuses instead on consensus within organizational subcultures; and a fragmentation perspective that focuses on "multiplicities of interpretation" (p. 107) and "transient and issue specific" consensus within an organization (p. 94). Though I did find evidence of subcultures and appreciate the importance of investigating the fluctuations and contradictions in an organizational culture that are the hallmarks of the fragmentation perspective, in this paper I focus my discussion of organizational culture on what I found to be dominant throughout AF.

Recruitment and selection

AF's recruitment and selection process is a critical part of the organization's overall "talent" strategy. As AF's growth accelerated in the summer of 2007, AFNS created a Vice President of Talent Development position that eventually developed into the role of Chief Talent Officer (CTO). Maia Heyck-Merlin, who occupied both roles, led "Team Talent," which was charged with conducting or coordinating network operations related to attracting, identifying, recruiting, selecting, retaining, developing, and evaluating teachers.[40] These efforts were aligned, carefully designed, and viewed as part of a coherent talent management system. Although they involved largely individualistic teaching quality strategies, they were widely understood to set the stage for many fundamental aspects of the organization's dominant culture and overall success by identifying and hiring people with a professed commitment to AF's mission, a growth orientation, strong critical thinking skills and a willingness to work collaboratively.

The strength of teachers new to the network is particularly important to AF as the network adds both schools to the network and grade levels to schools each year. This expansion means the numbers of teachers the network needs to hire has grown steadily each year and represents a significant proportion of the total workforce. In 2009-2010, AF employed approximately 370 teachers, roughly 42 percent of whom were new to AF that year (Curtis, 2011). Beth Meagher, AF's Director of Recruitment Outreach and Operations, reported that for the AF needed to hire just over 200 teachers for the following school year—91 in Connecticut

---

[40] Their vision was: "Each AF classroom and school will have awesomely talented educators working hard to make academic and character gains with their scholars. These educators, and their supporting team members in operations and network roles, will be excited to come to work each day because they will feel connected to our mission. Our talented Team & Family members will choose to stay at AF because we will be an environment that contains multiple career and learning opportunities and regularly recognizes excellence. As we build our team, we will invest wisely in infrastructure for the future and share broadly with our friends in education reform, " (NSVF, 2010).

and 113 in New York.[41] The recruitment team's efforts, a partnership with Teach For America (TFA), and hiring freezes in neighboring districts due to the economic downturn meant that AF had a deep applicant pool and was able to be selective in hiring, though they faced greater staffing challenges in particular geographies (like Bridgeport, Connecticut) and traditionally hard-to-staff areas like special education and secondary mathematics and science. In New York, they also faced competition for candidates from other charter schools and networks, including the Success Academies, Knowledge is Power Program (KIPP), and Uncommon Schools.

Recruitment and selection was handled in partnership between AFNS and the schools. The AFNS recruitment team ran outreach activities to attract a wide applicant pool, including e-newsletters and webinars, PD, and AF open houses. They helped to institute a summer residency program that employs early career teachers working elsewhere to teach in AF's summer academy and that connects strong candidates with the network. They screened applications and resumes, inviting strong aspirants for phone interviews with an AFNS recruiter. Those selected as finalists toured the employing school site and taught a demonstration lesson before debriefing it and interviewing with the recruiter, principal, and other school staff. If the principal wanted to hire the candidate, the recruiter checked the candidate's references before the principal extended an employment offer. As AF grew, the AFNS recruitment team created common tools that structured these activities, including detailed guides for the interviews, demo lesson observation and debriefs, and candidate evaluations. Their management of this process was intended to lift some of the burden associated with an intensive approach to recruitment from the principal; to ensure coherence in the selection standards across the network; and to minimize internal

---

[41] At the time of her interview in mid-May of 2010, Meagher reported that they had hired people to fill about 60 percent of those 200 slots. At that time, 57 percent of applicants to AF had been interviewed by phone, and of those interviewed, 26 and 33 percent participated in a finalist interviews in Connecticut and New York, respectively. Of those who completed finalist interviews, 41 percent had been hired in Connecticut, and 30 percent in New York. In other words, approximately 15 percent of applicants were eventually hired in both states.

competition for talent that could threaten collegiality across the network. However, because the principal made the final determinations about hiring and the candidate interviewed at the school, each party was able to assess the teacher's site-specific compatibility.

Each step of this common recruitment and selection process was designed to surface and assess the competencies that AF believes contribute to both individual and organizational effectiveness. AF's leaders identified eight broad categories they believe to be essential for successful practitioners within their network: breakthrough student achievement (results without excuses); effective and strategic planning; effective core instruction; data driven instruction; academic rigor; classroom and school culture; strength of character, and personal effectiveness. Within these categories are listed more detailed competencies, many of which were intended to assess the extent to which candidates' practices and "mindsets" are aligned with AF's mission, core values, and vision of excellent instruction. For example, within the category "breakthrough student achievement," each applicant was scrutinized for evidence that she or he "believes and expresses all students can achieve" and "believes in the AF mission of social justice." Within the "effective core instruction" category, AF evaluated the extent to which the candidate appeared to be "receptive to instructional feedback." "Strength of character" included the ways in which the candidate demonstrates orientations toward teamwork, professionalism, "whatever it takes," responsibility for outcomes, integrity, reflectiveness, and sense of humor (AF, 2009c). AF asked candidates for their undergraduate and graduate school (if applicable) grade point averages on their initial employment application, but expected that demonstrating strength in these many of competencies was likely to be associated with strong critical thinking skills and academic performance, yielding a group of teachers who were themselves successful students.

Network and school leaders explained that these core values and mindsets feature heavily in hiring decisions because they were believed to have broad implications for school and network success.[42] One network leader offered: "We're looking for people who are eminently coachable and who are hungry and want to get better, almost regardless of what prior experiences they have" (Interview15). Heyck-Merlin explained, "When we hire teachers we don't want classroom teachers, we want school teachers, people who feel responsible for the whole school. And the same with principals, you're not just responsible for your school; you're responsible for the network." Sara Keenan, AF's Director of Leadership Development, elaborated: "You might be a great teacher, but if your mindset and core values and collegiality and […] willingness to grow and get feedback aren't there, we're not going to hire you. And what that does is you just create these communities of people who are eager to learn, who are open to feedback, who are really enthusiastic."

The focus on teachers' mindsets and growth potential relative to their existing skills was one of several reasons that AF hired new teachers even though had not yet proven their effectiveness in the classroom. In the 2009-2010 school year, 77 percent of teachers hired had five or fewer years of teaching and 36 percent were first year teachers. Most of these new teachers came to AF through TFA, meaning that these candidates had already been screened by an organization that shared a similar mission and sought similar candidates for their program.[43] In the same year, 23 percent of the teachers hired were TFA corps members and another 33 percent were TFA alumni (Curtis, 2011).[44]

---

[42] Meagher reported during the previous two years, AFNS had been collecting data from this selection process so that eventually, they could combine it with evidence of teachers' effectiveness with student learning and learn more about how they might modify or weight the various criteria upon which they select teachers for the network.

[43] First year teachers went through a slightly different selection process since they do not have typically have formal teaching experience to draw upon.

[44] Since many of these new teachers had yet to be certified but were required to become so by New York and Connecticut, AF joined with KIPP, Uncommon Schools, and Hunter College to create TeacherU (now Relay

Doug McCurry, a former teacher and co-founder of Amistad and AF, who is now the

organization's Superintendent and co-CEO, offered several reasons AF hired unproven teachers

despite the organization's stated central concern for student achievement. First, while AF had

had success with new and "early career veteran teachers" who had been teaching for 3-6 years,

the organization struggled more with bringing veteran teachers into the fold.[45] McCurry

explained that it was often more difficult to entice veteran teachers to leave their districts and

accrued benefits. More importantly, while they have had veteran teachers thrive within AF, it

was often more difficult to teach them the "AF way." He continued, "It's been harder to bring

them into our culture, to [get them to] say, 'you know I've been doing it this way for fifteen

years, oh I'm suddenly going to do it a different way,' or 'I've been a classroom teacher where I

close my door and do my thing [and now] I'm a school teacher and really understand that there

are some things we're going to need to do as a school to make it work.'"[46] An academic dean at

one of AF's elementary schools offered an example of this commitment to collective practices:

> We say this in our schools all the time, I say to teachers in the hiring process,
> "You may have a brilliant lining up routine that you use in your classroom, you
> may have been honing it for years in your school wherever you are, but when you
> come here we think the value of every kindergartener lining up the same way

---

Graduate School of Education and independent from Hunter), a certification program based upon the approaches of these three CMOs.

[45] An elementary school principal added that in his experience, many of the young and relatively inexperienced teachers they hired through TFA or TFA-like programs were powerfully motivated by the belief that providing low-income and minority students with a high-quality education "is this vital work to make our country a freer and more just place," while teachers who had already taught for 3-7 years in traditional settings "are a little hesitant around [AF's] mission" and in their interviews, communicated "that they're not going to work 9 or 10 or 12 hours a day to make sure that our kids have all that they need" (Interview44).

[46] An AF principal echoed these sentiments when discussing why prior experience could be detrimental: "In many big places the experience you get is actually delivering curriculum and not approaching teaching in a very smart way, and so sometimes with experience you need to detox a lot from ways of thinking about it before. I think that actually the most important thing is somebody with very strong critical thinking skills because in that way you can really plan backwards, […] and have good judgment about what you're putting in and what you're putting out and why, and how to incorporate a lot of different information quickly and synthesize it to make sense for yourself. Otherwise what you get is pendulum swings from one curriculum to another, […] a lot of dogma without a lot of thought behind it. […] I think that experience without the critical thinking habits or skills [can be] extremely challenging and not productive" (Interview20). Reflecting the network's broader talent development strategy, she argued that because teaching can be so overwhelming, regardless of a teacher's prior experience, it was essential to help teachers to prioritize areas for their improvement and master them before taking on the next priority.

trumps you having this wonderful magical system that works for you. If that system is not going to work for everybody, that's probably not what we should do. […] There's such a value in a kindergartener knowing that no matter which of the adults in the building is lining them up it's going to be the same expectation; that's what we should do." (Interview39)[47]

Second, hiring new teachers was related to AF's effort be relevant to the broader school reform movement. "At some point traditional districts have to take rookie teachers in and figure out how to make them good, and so […] we need to take on that same challenge," McCurry offered.

Finally, he noted, AF's focus on hiring new teachers had an economic rationale:

You can probably hire two or three rookies for the cost of one twenty-year veteran, and at some point the economics are tough there. [One of the reasons AF is revamping its teacher career path model is because] the current teacher compensation model is broken. […] So for us to justify the twenty-year veteran who's going to make $95,000 a year, they better be a doggone all-star for us to justify them versus some else. […] I'm totally game to pay teachers more, I'd love to; I just want to make sure they're my best teachers.

This perspective represents a departure from more typical compensation models in which teachers move up a pay scale according to experience, course credits, and credentials.

Instructional framework and tools

AF's heavy reliance on what Wilson (2008) called "rare human capital," or those teachers with "exceptional educational backgrounds or unusual levels of commitment" (p. 17), creates obstacles for the replication and sustainability of their approach to teacher quality; this concentration of novice or early career teachers also poses challenges for network's ability to meet its goals for students. Both of these issues will be discussed toward the end of this essay. However, AF's deep commitment to their mission and the numbers of inexperienced teachers they hired together created strong incentives for AF to develop the resources that might support

---

[47] This dean added that if the shared approach or procedure was not working after a trial period, "we're not going to stick with something poor just for the point of having it," so the school would revisit and change it as a group.

novices in achieving at least a minimum standard of practice across their classrooms and schools, and then in growing beyond that. As one elementary school principal explained,

> Our teachers are really driven. We hire for that and they're amazing, so I have to say 80 percent of what happens is their own reflective nature, their ambition about hitting their goals. But it does take more than that. I mean, when we opened the school last year 10 out of our 14 teachers were first year teachers. […] And so to hit goals with that team there has to be some sort of critical support. (Interview18)

This "critical support" took the form of a common and increasingly coherent set of instructional tools and ongoing supports for professional learning and practice, which supported but were also animated by, this staff.

The Cycle of Effective Instruction and the Essentials of Effective Instruction

AF codified its expectations for teachers' core instructional responsibilities in two documents—The Cycle of Highly Effective Teaching (see Appendix C), and the Essentials of Effective Instruction (see Appendix D). The Cycle specified the key resources that AF teachers were to draw upon as they set ambitious and measureable goals for student outcomes for the school year and each interim assessment (IA) cycle, and as they crafted the long-term plans that would help them to achieve these goals. Then, the Cycle set expectations that teachers would: (1) create and modify unit plans; (2) plan lessons; (3) engage in instruction, including ongoing assessment of student progress and interventions for struggling students; and (4) administer IAs and use the results to modify goals and instructional plans. At the center of the Cycle was a reminder that these processes should also draw upon or contribute to AF's core values; constant learning; personal organization and effectiveness; and student and family investment.

The Essentials framework provided additional detail for the third step in this overall cycle and was seen as one of the central tools for teaching quality support throughout the network.

Maia Heyck-Merlin, who, in the spring of 2010 had recently transitioned into the role of Chief

Operating Officer (COO), reported that as the network grew AF's leaders looked for ways to

more systematically support teacher quality. "[W]e came down to these two things: define the

picture of excellence […], and then help people to get there." The Essentials were drafted in

2008 as a way of explicitly articulating a shared vision of excellence in teaching, with AFNS

leaders drawing upon their own experiences, similar work by peer organizations, and others'

research on teaching to identify practices they believed to be most effective in raising student

achievement across grades and subject areas.

The 2009-2010 version of the Essentials included ten main categories and 24 elements

that fit within them. It was designed to be straightforward and parsimonious to support classroom

use. McCurry's lengthy description of a strong AF teacher offers a useful narrative illustration of

the Cycle and the Essentials as the codification this shared vision for instructional excellence:

> [W]e're Achievement First, right? So demonstrated breakthrough student
> achievement gains are one key thing. Did your students show tremendous growth
> and did they meet our own internal benchmarks and the big goals for the school
> and for your subject area? At the end of the day that's the most important thing.
> [And then] there is a special sauce that goes beyond what would be on paper, but
> if I were to look at teachers that are great, first of all they're extraordinarily well
> planned, […] they've actually back mapped the year into clear units, then they
> back mapped each unit into what the assessment is for the unit, and it's a rigorous
> assessment that really makes kids think. And then they've thought about what the
> enduring understandings, essential questions for that unit are, then they've back
> mapped that into clear bite size aims and are very well planned. And then when
> you drill into that then they're really well planned for each day […]. You would
> walk into their class and you would see a super clear bite-sized, measureable,
> rigorous aim that was driving instruction; you'd see some way at the end of class,
> typically by an exit ticket […] where they're measuring that. And then you would
> see a high degree of mastery of the day's aim pretty consistently, or if not they
> would really be using that data to inform other [aims]. You'd see them being
> super, super clear about the instruction, using economy of language to figure out
> the most effective and efficient way of explaining things to kids; you'd see tons of
> what we call "at bats," so whatever the aim is the student's getting tons of practice
> in it. […] I think sometimes people think of that as just slogging away [but] it
> doesn't have to be that […]. And then really clear expectations about the quality

of work that they would produce, and then super clear expectations about how they're going to discuss that work, and do they have the ability to kind of riff off their classmates to do that. […] Within that, the students are doing the heavy lifting, they're doing the majority of work. The teachers really preplan their questioning, they're super clear about exactly what student oral response excellence looks like, what student written response [is like]. They're maniacal about time and being really efficient and super well-paced, making sure a ton's happening. And…they have the structures and routines and expectations in class that are super smooth and don't take time; 100% of the kids are on task at all times, and then 100% of the kids are engaged. […] The best teachers have had some sort of hook to get the kids really into it and they've used some pretty subtle but effective techniques throughout the lesson to make sure that 100% of the kids [are engaged], so if they ask a question they use a stand up, sit down, use a think pair share, use mini white boards, show me on your hands where you are, so every single kid is engaged. I think it's all of it, and then all those things put together in a really seamless way is what really strong teaching would look like.

In discussing what it meant to be a good teacher in AF, regional superintendent Marc Michaelson echoed several of the comments made by people discussing the qualities AF looked for in teachers. He noted that in addition to supporting high-levels of student achievement and their character development,

> I think the thing that makes, maybe a great Achievement First teacher different than a great teacher in another public school system is that we have very much a team approach in our schools, and so you're not just a lone ranger closing your door and being terrific in your own way […]. I think there are a lot of things that we do together as grade level teams, there are things we do in curricular teams, there are things that we do Achievement First-wide […] and so I think there is some level of kind of team that is infused into being a great teacher as well.

When asked whether or not the teachers each had described were regularly found in AF's classrooms, both responded in similar ways. McCurry observed that while most teachers in the network were "pretty solid" on the classroom management and the basic lesson cycle elements of the Essentials, in "very few" classrooms did all of the things he described above take place on a daily basis. Michaelson suggested that AF's coaching model, which will later be described in greater depth, was key priority for the network as they sought to move people from "good to very good and to exceptional."

As the articulation of the network's evolving understanding of what it needed to require of their teachers, and as a set of principles and practices for teachers and instructional leaders to draw upon in their work, the Cycle and Essentials provided educators with a framework around which to organize.[48] The Essentials structured many of the other teaching quality management efforts throughout the network, including teacher recruitment, selection, and evaluation; the summer, network-wide new staff induction and training; and many aspects of AF's ongoing professional development efforts (for more about this work, see also the "Coaching" section of this essay). While building shared understanding and practices around the Essentials took a great deal of continuing work, AF network and school leaders overwhelmingly reported that alignment of organizational activity around the Essentials had made AF, in Heyck-Merlin's words, "a much more cohesive organization, [with everyone] working towards the same vision."[49]

Curricular tools: Scope and sequence, interim assessments, Athena, and other resources

The first steps listed on AF's Cycle of Effective Instruction were the creation of AF's scope and sequence and their interim assessments (IAs). IAs were a set of grade and subject specific

---

[48] The Essentials were meant to provide structured guidance for instruction in AF, but were not intended to be used as rigid rules. Although teachers are introduced to the Essentials during new-staff training, Alex Freidus, Director of Teacher Leadership Development, noted that some people needed additional support to learn to use them in the ways that AF intended. Some teachers, especially those new to AF, sometimes "look at the Essentials of Instruction and see it as a checklist instead of as a set of principles that can guide decisions you're making," Freidus explained. This can be problematic because when they see the Essentials this way, they may focus too much attention on ticking off categories on the list rather than engaging in deeper thought about the lesson's purpose (e.g., Do I have I/we/you portions of the lesson, instead of asking, what is the purpose of the mini-lesson relative to my instructional goals in the lesson?). Viewing the Essentials as a checklist can also frustrate teachers who may then understand the Essentials to sometimes constrain their ability to design effective lessons. Freidus suggested that these teachers and their coaches often needed support in thinking about how to use the Essentials in principled yet flexible ways.

[49] In addition to working on within-school alignment around the Essentials, AF also worked to build cross-school Essentials alignment. For example, in addition to regular work across schools through cohort meetings and the coaching of principals, McCurry and the regional superintendents worked to establish shared norms about the Essentials among network and school leaders by creating an Essentials rubric; watching instructional video together using it; and doing formal co-observations of teachers' instruction with other network personnel and all of the principals and academic deans. The co-observation process also provided a snapshot of instructional quality across the network at the beginning and end of the year; provided leaders with an opportunity to discuss, prioritize, and plan for support and development; fostered discussion about whether or not the Essentials adequately captured what stood out as important in the co-observations.

benchmark assessments administered five times each year, or roughly every six weeks, in reading, mathematics, and writing, and less often in social studies and science in the middle grades. The instructional scope and sequence detailed the teaching and learning objectives aligned with state standards that were taught within each IA cycle to prepare students for the state assessment, promotion to the next grade, and eventually, college. In many ways, then, the IAs based on the state standards may be understood as the core of AF's infrastructure of practice.

In 2009-2010, AF had developed a common scope and sequence for math, English language arts, and history, with a new K-4 writing scope and sequence. Because AF operates in both New York and Connecticut, they developed two somewhat different versions of these tools so they were aligned with the states' respective standards and assessments for grades assessed by state exams. Both the IAs and the scope and sequence were developed by AFNS' curriculum and professional development team, often in collaboration with educators from across the network. They were also circulated to all teachers throughout the network for feedback and editing. Having the teachers receive the IAs before each cycle was seen as important within the network because the IAs further articulated the learning expectations outlined by the objectives and standards in the scope and sequence.

Athena, a custom-built online assessment platform used by AF since the fall of 2007 provided all of AF's teachers and leaders with almost immediate access to network-wide student performance data and tools for their analysis. To facilitate a timely turnaround of the data, AF's schools were equipped with production quality scanners that could, at a very rapid speed, scan bubbled answer sheets as well as entire test booklets used for all grades and subjects K-4, and in reading, writing, and history in grades 5-12 capture the text of the test as well and written student responses. These data were funneled directly into the Athena platform.

Athena was intended to give teachers information on their students' progress relative to their goals after each IA cycle, and to do so in ways that support subsequent instructional planning. Ultimately, Chief Information Officer Harris Ferrell reported that the goal was to provide teachers with "the best information we can so that they can maximize the impact of their instructional minutes," deciding what to review or reteach, and how to organize students to do so. Second, Athena was meant to support teacher self-reflection by providing detailed feedback on their effectiveness. However, unlike student outcomes on state standardized tests, Athena was not intended to be an evaluative tool. Ferrell explained,

> By taking the evaluation out of it, teachers don't feel like they have to game the system to try to produce a result that's inconsistent with their students' learning. […] They can say, "This is giving me an honest picture of where my students are and where their gaps and their strengths are, I can be a better teacher with that."

If teachers feared losing their jobs as a result of their students' IA performance, Ferrell worried, "then you're going to approach your instruction differently, you're going to approach the use of that data differently, and you're really going to game the system, and everyone loses."

Athena's platform was designed to support teachers' use of the IA data in these formative ways by allowing them to easily review and organize complex data. For example, within Athena, teachers could examine student scores by standard or by question, and by grade, class subject or section, or by individual student. They could also engage in error analysis, looking at the distribution of student mistakes by test item and easily accessing embedded links to the items, along with student work or written responses scanned into the system for every open-ended question. Reports showing student performance for a single IA cycle or their growth over time were also available. Standardized color-coding of the scores and sorting functions allowed greater visualization of these data. What is more, teachers were able to access IA results from across the network, allowing them to assess how their school, grade level, and students were

performing relative to their peers. If they struggled in a particular area during an IA cycle, they could reach out to teachers at other schools who met with greater success.

To help teachers and instructional leaders to learn to approach data analysis in productive ways, AFNS developed a series of trainings and protocols to guide this work. During new staff training, teachers received training and a written guide on how to use Athena in a technical sense. The network also offered teachers optional network-wide professional development about more in-depth data analysis twice a year. Most importantly, at the beginning of each IA cycle, each school hosted a student-free "data day" during which the staff convened to engage in collaborative data analysis and instructional planning (for more about data days and opportunities for teacher learning on these occasions, please see the section "Data analysis and planning" later in this essay). In addition to IA data, teachers were encouraged to use other sources of evidence about their students and student learning including classroom observations and a host of other assessment data that varied some by school and grade. For example, the Fountas and Pinnell Benchmark Assessment System was used in grades K-8 to assess and guide reading progress, and these results were expected to be used in planning. Some of the elementary schools had adopted the USI's STEP system for literacy as well.[50]

Much of teachers' time on data days was directed at creating a Data Driven Plan (DDP), or "tool for developing [teachers'] hypotheses about the biggest challenges facing [their] class, and the instructional steps [they] will take to respond to them" (AF, 2009d). The DDP was document that structured teachers' assessment of classroom and student progress relative to established achievement goals; efforts to prioritize and plan for any whole-class re-teaching or

---

[50] At the end of the year, TerraNova mathematics assessments were administered in grades K-2, and all 1-12 grade AF students also took the Developmental Reading Assessment (DRA) or Degrees of Reading Power (DRP) upon entering AF and at the end of the school year. Results from state standardized assessments and other standardized tests like the SATs or Advanced Placement exams were also collected and available for planning for the following year as well; Like the F&P and STEP results, these assessment results could also be compared across regions.

review; and their identification and planning for small group or individual student interventions inside or outside of the classroom. Correspondingly, the DDP called for teachers to plan concrete next steps based on these analyses, like updating their goals, aims sequence or long-term plans, and creating lesson plans based on these resources and the AF scope and sequence. Teachers also received a detailed DDP Guide organized into three main steps of "effective data driven planning": identifying big needs, diagnosing specific issues, and planning instruction. The guide contained probing questions and lists of tools (e.g., specific reports from Athena, or a copy of the scope and sequence) to support their thoughtful response, for each of these steps. It also includes a "Checklist of Analysis and Planning" to assist teachers in completing the DDP process; a list of common DDP pitfalls to avoid; and two illustrated examples of the DDP process and products.

Besides supporting teachers in their work, Athena was intended to help school and network leaders in their labors. For the principal, Athena provided clear and timely overviews of how each grade level was performing in each subject. Ferrell explained, "[T]he principal's role is [to look a these data and say] okay, where do I need to prioritize my time or my organizational resources based on the outcomes I'm seeing holistically across my school?" Similarly, Athena allowed the superintendent and regional superintendents to get an overall picture of achievement in the network and to target AF's resources appropriately. Looking at data across the network, they could also identify areas for improvement that might benefit from a network-wide focus. Shelley Thomas, Product Manager of Athena, explained AFNS leaders reviewed data from Athena to ensure that the IAs were indeed predictive of success on state assessments and of students' success in subsequent grades. If they noticed a particular trend, for example, if "there seems to be a real challenge consistently in New York middle school reading, […] we'll get the

appropriate parties together and think about what kind of intervention do we need […] at a school level to get things back on track."

In addition to these shared curricular tools, there were other non-negotiable aspects of AF's model, including having an extended school day and year to allow for more instructional time, and the use of the Fountas and Pinnell Benchmark Assessment system (F&P) for guided reading. Yet different schools organized their instructional time differently, organized teachers' work somewhat differently (e.g., departmentalization in elementary school subjects, or not), and adopted different curricula or supplementary assessment programs to guide work in particular content areas (e.g. Investigations versus enVisionMATH in mathematics; some schools' adoption of the UEI's STEP system or Lucy Calkin's writer's and reader's workshop models for structuring literacy time in addition to F&P for guided reading; FOSS materials in some schools for elementary science instruction, etc.). The potential tension between a highly structured set of shared curricular tools and a great deal of school-level flexibility in more detailed curricular choices appeared to involve ongoing dialogues between AFNS and the schools, largely negotiated through strong professional relationships established through the network's design and formal and informal instructional leadership roles.

Instructional leadership and organizational design to support learning and improvement

This shared set of hiring practices and instructional tools provided AF with key elements of an infrastructure of practice, including the resources around which a coherent set of values and instruments that guided the content and method of teachers' work; standards for assessing practice; and a framework to organize and codify many of the network's improvement efforts. In AF, formal leadership roles and organizational arrangements within and across schools added to

AF's infrastructure of practice by helping to establish strong professional relationships that provided teachers with important opportunities to access and learn to use these instructional resources, to share their learning with others, and to improve upon these tools. The organization of leadership also helped to protect the core work of instruction and instructional leadership by distributing responsibility across roles and parts of the network.

Below, I first provide a description of key network and school leadership roles related to instructional quality support and management, and how these positions are organized within AF. Then, I highlight four critical aspects of the instructional leadership work at AF that drew upon and contributed to AF's infrastructure of practice and the development of individual practitioners. These included informal mentorship; coaching; data analysis and planning; and formal management of school-based talent.

School and network leadership and organization

At the school level, the principal's role as an instructional leader was partially sheltered by the presence of a director of operations who assumed responsibility for many of the school's non-instructional functions, and a dean of students, who focused on maintaining school culture and student behavior. AF schools also had a school support team, which generally includes at least one learning specialist who coordinated interventions for students within the school (special education and others) and a school social worker. The AF elementary school model also called for least one academic dean who, together with the principal, provided teachers with targeted PD.[51] This included weekly PD sessions on Friday afternoons when students were released two hours early as well as weekly or biweekly individualized instructional coaching for every

---

[51] In part because of differences in the funding of charter schools in the two states in which AF operates, for example, elementary schools in Connecticut tended to have a single academic dean while those in New York generally had two.

teacher. Principals also attempted to organize their work so they were always teaching children in some way, by doing something like pulling a small literacy intervention group for one or more six-week stretch or co-teaching a mathematics block with a new or struggling teacher for several months. They reported that this helped them to stay connected to the realities of instruction, get to know children in their schools as students, model instructional strategies, and do work that they frequently professed to enjoy and miss.

Other important school leadership roles included trained grade level or content area chairs, selected in part for the strength of their instruction and students' results and partly for their skill or potential as leaders of adults. These chairs were expected to establish performance-oriented teams that met once or twice weekly to plan for instruction and discuss student progress. School schedules were deliberately arranged so that grade level or content area teams shared some of their 90 to 120 minutes of planning time during the school day. Chairs also led their teams on data days at the beginning of each IA cycle (see the "Data analysis and planning" section of this essay for more about this work). In elementary schools, two teachers were assigned to each Kindergarten through second grade class of approximately 30 students so that there could be more flexibility for organizing young students in small groups for targeted instruction within their classrooms. These pairings were made strategically so that more advanced teachers were placed with new teachers or others who were likely to need more consistent mentorship and support.[52]

---

[52] Outside of these general outlines, there was a good deal of variability in the leadership structure at the school depending on the strengths and needs of the faculty. For example, the principal of one AF's network had carefully built a school "cabinet" consisting of the school leadership team and grade level chairs. This cabinet shared significant responsibilities in school planning and decision-making, representing their respective grade levels or content areas, as well as helping to create buy-in with the teachers they led for changes or new initiatives within the school. But because several key members of that cabinet were transitioning to other jobs within or outside of the network, he planned to dramatically change the leadership structure at the school the following year to allow less experienced teachers and deans to ease into new leadership roles and responsibilities.

Network-level leadership roles and structures facilitated teacher and school leader learning within and across schools. AFNS provides training for teachers in the summer, regional content area trainings for middle schools, and two AF-wide PD days in the fall and spring. Other AFNS teams, such as the curriculum and professional development team, regularly provided trainings and consulted with schools, providing additional support tailored to teams' needs. School leaders interact with peers from other schools in regular training and cohort meetings.

AF's top leadership was split between two founders and co-CEOs, Dacia Toll and Doug McCurry. Toll acted as the organization's president, leading AF's operations work, including the teams responsible for finance, facilities, recruitment, and leadership development, as well as the network's external relations and marketing teams. As Superintendent, McCurry led those responsible for work more directly related to school support and academics, including the curriculum and professional development team, special education team, and groups focused on data and information technology. Furthermore, he oversaw AF's three regional superintendents. McCurry and the regional superintendents each coached a small portfolio of principals, often spending considerable amounts of time each week helping them to problem solve and connecting them with resources within and outside the network. Despite this division of labor and their myriad other responsibilities, McCurry and Toll, as well as other senior leaders at AFNS, worked to stay close to the work of the schools and to instruction. COO Heyck-Merlin offered:

> I think one of the things that's really interesting to me about AF and that I'm really proud of is that even the folks who are most senior still have a heavy, heavy close […] impact on day to day teaching. So it makes me really proud and I think signals what we're all about, that Doug and Dacia themselves teach teachers how to lesson plan on day one of new staff training, that all of our principals teach every single day, that all members of our senior management team […] have been teachers.

This attention to instruction and the day-to-day operations of the schools, together with the organization and design of the AF network and its leadership at the network and school levels, were intended to cultivate a sense of shared responsibility for student performance across the network. Still, managing the distribution of responsibility and power among AFNS leaders, school leaders, and teachers required consistent organizational time and attention since establishing a workable balance was likely to have serious repercussions for the organization.

For example, while the charter management agreements (CMAs) that bound schools and AFNS together helped to outline the respective responsibilities, as AF grew the network's leaders found it important to be much more explicit about the AFNS-school relationship. During the 2009-2010 school year, the principal cohort worked with Superintendent McCurry and several other senior AFNS staffers to make more explicit AF expectations about the balance between school autonomy and network strengths. These deliberations resulted in a draft document entitled "Achievement First Shared Beliefs and Practices" (AF, 2010). In the opening section of the Shared Practices draft shared with me in the spring of 2010, the authors restated their shared commitment to AF's mission, writing that to meet these goals, "we all understand both that principals will need wide and deep decision rights in order to drive achievement at their schools and that there will need to be a set of practices shared by all AF schools in order to provide the support necessary to aid and sustain schools in their work," and that "this document endeavors to capture this sweet spot of power to lead and power of the network" (p. 1) After establishing a set of principles to guide relevant considerations about these questions, the document contains a table organizing AF's practices or areas of practice into three different categories: (a) consistent practices, or those that were expected to be consistent across schools (both the what, and the how, of these practices had to be the same, though principals had input

into the how); (b) consistent outcomes, or those that were expected to have a practice toward some end that met some minimum criteria (what), but the "how" was left up to the school teams, though AF aimed "to provide strong exemplars and best practices for principals to use and modify" (p. 7); and (c) best practices, the use of which were left to principal and school team discretion. The table included rationales to explain why practices were categorized as they were.

This Shared Practices document provides a useful artifact of AF's leaders' thinking about these issues as well as an example of the collaborative process through which they were negotiated. The draft of this document also illuminates several of the ways in which some of the questions about collective work intersect with the value of individual agency in complex work like school management in AF. AF had come to see some degree of standardization across the network as necessary to monitor and motivate performance, to learn from the network's work, and organize to support leaders and practitioners. Still, AF also saw innovation and the tailoring of aspects of AF's model given the particular contexts of educators' work as essential for the organization's success and development. Similarly, while AF valued people who were willing to work in teams and agreed to do some things the "AF way," the organization also valued people who demonstrated unusual initiative and personal effectiveness. Retaining these individuals over time is likely to require that they have some room to innovate in their work and to contribute to the organization's development as well.

Even with the balance between consistency and flexibility for schools across the network more clearly articulated through the Shared Practices document draft, AF still expected that strong professional relationships characterized by trust would continue to be a central resource for negotiating some of these complexities. McCurry explained,

> No one goes back to documents as they make decisions, and so we really want
> this to be a relationship where the principal feels like, "I really can do what I need

to do to make this school great," but then says, "I have a responsibility to be really in deep dialogue with my assistant sup, because they've done this before and they're pretty good at this, and I want to make sure that I'm heading in the right direction."

Regular meetings between the superintendents and principals were intended to facilitate this kind of relationship and dialogue.

Although the balance between what was consistent and flexible was continually evolving and some people worried AF was not yet striking the right equilibrium, the elementary school leaders I interviewed uniformly expressed enormous respect for AFNS and the people working there, understood the rationales behind some degree of standardization across schools, and appreciated the way AFNS engaged in making decisions about what needed to be "shared." As an academic dean explained,

> The network is very good about leaving space for healthy debate, for asking for feedback on things, for letting schools and principals and everyone, hash things out. I don't feel like these are top-down directives, I feel like, "This is our best thinking on what would be best for everybody, and now let's talk about it." I think that the fact that there's room for that voice makes it feel okay with me. (Interview39)

A principal at a different school described his experience working within with AFNS as overwhelmingly positive, even if it occasionally involved compromise, because he trusted his colleagues there to work from similar principles to pursue their shared mission:

> It really is my belief that even as things are standardized, even if that occasionally means that I may not have, or our school may not have, the final say on what an interim assessment looks like or the direction that a particular program is going academically, I have total faith that my teammates at network support who are making those decisions have scrutinized the data the same way that I would and have and done research in the same way that I would and are making decisions based on the same values that I would [use to] make them. And even if I might have arrived at a different answer, I always think that the one that they've arrived at is probably as good or better because of that. I think that's just a trust thing.

While he reported being aware of some tensions around AFNS-school relationships, he reported that he worried more about a "redundancies and inefficiencies in some of the ways that some of the AF teams [operate]" because of a reluctance to standardize practices earlier (Interview44).[53]

Like the framework created by AF's instructional tools—and especially the agreement on common aims and their measurement throughout the network—strong professional relationships among staff members with shared commitments to collective work, and to student and adult learning, contributed to AF leaders' ability to embrace considerable flexibility at the school level. AF's infrastructure of practice also appeared to support network and school leaders in avoiding a strong bureaucratic leadership model based upon top-down mandates, rules and regulations, and to simultaneously avoid situations in which each teacher and school is expected to learn how to be successful in their work on their own. Instead, AF seemed to make use of and foster the deepened sense of accountability, responsibility, and loyalty that can come from feeling ownership over one's work within parameters set by guidance and oversight from AF.[54] These arrangements also meant that AF was poised to learn from the "experiments" taking place across classrooms and schools. Below, I offer brief descriptions of four critical strands of

---

[53] He elaborated: "There are [so many] different academic deans who are coaching teachers to become better, and at the same time there's somebody from AF network support who's working on making a new curriculum for the same program. […] Is there a chance that somebody's investing a lot of effort and they don't need to because somebody else has already built out that subject? […] So if anything I think the thing that would make it better is if we did move quicker in establishing best practices, if we did more clearly say in more subjects and more decisions that there is really a best practice that has emerged, at least for now, and that we should all do it or try to do it in the same or in a similar fashion" (Principal Interview44).

[54] Rowan (1990) developed two models for the organizational design of schools, which he named "commitment" and "control" models. He wrote that these two strategies "rely on different organization design features and attempt to affect different school processes to achieve school effectiveness." He continued, "The control strategy involves the development of an elaborate system of input, behavior, and output controls designed to regulate classroom teaching and standardize student opportunities for learning, and the expected result is an increase in student achievement. The commitment strategy, by contrast, rejects bureaucratic controls as a mode of school improvement and instead seeks to develop innovative working arrangements that support teachers' decision-making and increase teachers' engagement in the tasks of teaching. The assumption of this approach is that collaborative and participative management practices will unleash the energy and expertise of committed teachers and thereby lead to improved student learning" (p. 354). One way to think about AF's approach to instructional support and management is as a novel blend of these two models via AF's infrastructure of practice.

instructional leadership work that provide additional examples of the ways in which this

approach to management, support, and learning played out within the network.

Informal mentoring. Interviewed teachers and school leaders reported that informal

mentoring among school-based personnel was common and was an important source for learning

about teaching. As one veteran elementary school teacher explained,

> I think that the thing about [this school], and probably AF across the board, […] you're getting tons of help, but you just work your tail off. You work so hard. And I was very lucky. And I think that if you ask pretty much anyone in this building they'll probably all say the same thing, I was very lucky because so-and-so taught me. There's very much a culture of, there's a few people that you work with, [who] teach what you need to know […].

> Rosenberg: Somebody just said that about you today.

> I: That's so funny. Because I feel that way about [a woman], who now is [working for AFNS]. She is a phenomenal teacher, and I came in and she taught me everything I needed to know. And I asked her questions and she was patient and she gave me the help I needed and then she moved on to other things and I kind of stepped into that role and I've learned it. Of course you have to do a ton of learning on your own, because […] they can't tell you everything. You have to ask a lot of questions of a lot of people, but you have chances to observe and […] we spend a lot of time together and you learn. And then it's kind of like you passing down the skill. (Interview41)

This description of informal mentorship as a facet of AF's culture or ethos instead of just being

attributed to a few isolated individuals was echoed across teacher, network, and school leader

interviews. In addition to seeking support from colleagues in their schools, teachers reported

reaching out to their peers at different schools on the basis of recommendations from their

coaches or because of their impressive IA results. Principals and deans reported making liberal

use of their cohorts or a specific informal mentor at their school or AFNS, and AFNS personnel

frequently discussed essential informal support and collaboration from colleagues.

Coaching. Every teacher and school leader had a coach who was charged with engaging

them in continuous improvement of their work. Within AF, the coaching of teachers was widely

understood to play a critical role in the organization's ability to succeed, particularly as it grew, and as such was expected to be the core work of school leaders. However, 2009-2010 was only the second year that every teacher within the network was paired with a school-based coach—either a principal, principal-in-residence, academic dean, or sometimes other teacher leaders—who were themselves selected for their roles in part based on their own success in the classroom.[55] As a result, coaches' roles were still evolving. However, at the center of the model was the expectation that coaches would help the teachers in their portfolio (ideally no more than 8-12 teachers) to improve in their practice regardless of where they were in their development as practitioners. Coaching was also seen as a way to recognize excellence among teachers, as a recruitment tool with potential to attract applicants with the growth orientations that AF prizes, and as a tool to reinforce those mindsets in existing staff. In fact, almost all of the teachers I interviewed spoke highly of their experiences with coaching precisely because of the assistance the formal relationship had given them in improving their teaching.

AFNS leaders reported that previously, coaching or mentoring within AF had been dominated by a lesson observation and feedback cycle. In 2009-2010, AF was deliberately trying to establish a different role for coaches so they would be more effective in moving teacher practice and student outcomes forward. Keenan explained that beginning in the second year of the new network-wide coaching model, AF had begun to focus on helping coaches learn about "coaching teachers to mastery on a sustained learning goal." For example, for a newer teacher, an initial learning goal might focus on "classroom routines and procedures, to make the first ten minutes of your class super tight." In this case, coaching to mastery might involve something like the following:

---

[55] Teacher coaches were typically given smaller portfolios of between 1-3 teachers, in addition to a stipend and release time from their classroom work, to facilitate this work.

We're actually going to make a teacher-learning plan […] so that that goal is at the top. So instead of a check-in, where you're going through all these topics, the goal is at the top and you think, I'm not sure how long it's going to take us to do this, maybe three weeks, maybe four, but we're going to talk about, "What would it look like if it was really excellent?" and we write that down. […] And then you spend some time at the beginning really clarifying that picture, so that might mean let's go look at a teacher who's amazing at this, let me model, let's look at video that we have. So you have some way of really capturing what you're talking about, not just saying it, but seeing it. And then […] we're going to sit down and plan together what your routines are for the first ten minutes, and then we're going to plan, how are you going to re-teach [the routines to the class], because we've kind of lost it. And then I might go in and observe you and then we're going to talk about that. I might come in and model for you. […] And then you move on to a new learning goal once you see progress, growth, mastery in that.

For more advanced teachers, Keenan suggested, "a lot of their learning goals are around increasing rigor, [or] integrating Understanding By Design[56] units into [their] ELA class, things like that. So they might not meet quite as often, but we're really doing intensive geeking on instruction, co-planning."

McCurry explained that they had adopted this approach in part because "we want to avoid *merely* subject specific feedback" [emphasis mine]. He suggested that instead of telling a teacher that, for example, they had delivered a lesson on fractions in a confusing way, as a coach it was important to step back and say something like,

"How do I get you better fractions knowledge?" But […] it's actually in some ways easier to give that feedback sometimes than it is to […] step back and say, […] "There's a fundamental problem that […] you're increasing the rigor [of your lessons as you move from modeling to guided practice to independent practice] while decreasing the scaffolding and that's creating all sorts of problems in your classroom." And we want to make sure we hit that, which is Essentials focused, as well as the content.

In other words, McCurry continued, "We want to get away from, "'I watched this one lesson and this is how this one lesson could get better' [and move to] 'I watched this lesson, […] and I've

---

[56] *Understanding by Design* may refer to a book by Grant Wiggins and Jay McTighe, first published in 1998, and the framework for curricular and instructional planning through a process called "backward design" that they laid out, emphasizes establishing a curricular unit's goals and assessments before planning the lessons, tasks, and activities that will help students to succeed with them.

looked at your data, and here are the one or two things […] that you need to get better on in general to become a stronger teacher.'" He went on to explain, "That's why we think the Essentials are so important, because it grounds [the coaching process and the content focus]: How do you get better at that Essential, versus how do you fix that lesson?"

Indeed, although "knowledge capture and sharing" was a major priority within AF, with leaders and educators throughout the network working doggedly to develop improved systems for sharing information—like building systems for sharing unit and lesson plans and other instructional materials aligned with particular content standards that some of their best teachers have developed—coaches were still viewed as a fundamental resource for supporting teacher learning and practice. Ferrell explained,

> [I used to have] this vision that you take these interim assessments, and it's going to show you how students are performing. You find a classroom where the students are doing the best, you go to that teacher, say give me your lesson plan from when you taught this. Your students did better on this than anyone else, give me your lesson plan, I'm going to put it in this database and I'm going to link it to that standard and we've got the performance against that standard. Next time a classroom takes the assessment and they've got low performance on the standard, they click on the resource link and up is going to come your lesson plan, and now they're going to be a better teacher with that lesson plan.

Over time, however, Ferrell reports, his understanding of how AF might work to support strong teaching across the network evolved as he began to realize that the lesson plan itself may not have been the driver for successful student performance, at least not by itself; other features of the teacher's instruction might have been more important. In addition, a struggling teacher might need more than even excellent plans to help them to become more effective. Instead, he suggested, continued access to a coach who could help to diagnose both student and teacher weaknesses and help them to engage with appropriate supports is what was necessary:

> What I really require is a coach sitting down with me, understanding why I'm struggling as a teacher, or where my weakness are as a teacher, who can then look

at those five lesson plans aligned to that standard and say this is the one, [due to] a couple things: One, have we gotten a clear diagnosis of why your students have struggled on that standard, is it operational, have they selected the wrong facts from the problems, do they understand the structure of the equation they're supposed to build? What is it that they haven't understood correctly to perform well in the standard? And then if you understand that then you can go and look at great lesson plans that help [with] perimeters and what not. But if you haven't really gotten a clear understanding of the diagnosis then just lesson plans aligned to standards isn't going to help. So that's sort of piece one. Piece two is [if I'm a struggling teacher,] my challenge might not be my lesson plan, it might be my classroom management, or it might be that my mini lesson takes 35 minutes and I only have 10 minutes of independent practice instead of 10 minutes of mini lesson, 35 minutes of independent practice, those types of things. You really need someone who really understands where I am as a teacher to help me understand when I look at my data what my challenges are, how I would take those materials, use them to the highest effect for my students' learning. I think that coaching piece is critical to it and I don't currently know of a system that can play the role of that coach. So one of the things we're thinking about is how do we endow our coaches with knowledge of great materials so that they can apply that knowledge of those materials to the specific context of me as a teacher, and my own professional development as a teacher.

Ferrell's comments highlight the potential of the coaching model for supporting teachers in making good use of the resources and structures of the network in the particular context of their development and work.

AF aimed to support coaches in doing this complex instructional work in several other key ways. At each school, one of the principal's central responsibilities was providing support and coaching to the other coaches in the school. However, AFNS also played a role in supporting coaches in learning to do their work. For example, McCurry reported that one of the chief roles of the Superintendent and regional superintendents was ensuring "the principal and all the deans are really strong coaches of teachers." The superintendents also sought to provide strong models of coaching for the principals in their portfolio. In addition, network-wide workshops in July and November for coaches, and designated coaching support time during academic dean and principal cohort meetings, assisted coaches in learning to do their work.

152

These PD sessions introduced teachers to a variety of tools that AFNS had developed to support them in their work. Most importantly, AFNS drafted an evolving tool called "Coaching the Essentials," which broke the Essentials into components and established a rough order for their mastery by a teacher, moving from basic management to core instruction to more "advanced" topics such as high engagement strategies, rigorous instructional strategies, and high investment and character development (AF, 2010b). Next to each of these components were listed coaching methods and strategies that may be useful in supporting teachers in that topic and a corresponding list of hyper-linked resources that were available on the shared AF server, including, for example, internally developed resources like annotated lesson plans or checklists; relevant video from the Lemov Taxonomy; excerpts from Canter's (2001) *Assertive Discipline*, Saphier and Gower's (1997) *The Skillful Teacher*, Wiggins & McTighe's (2005) *Understanding by Design*, and TFA's Institute curriculum; and research studies. The coaches I spoke with highly valued this tool and their ability to use it flexibly with the teachers they coached, and a senior leader within AF applauded the tool's potential for creating, capturing, and disseminating a common vocabulary for instruction and instructional coaching.

Despite all of this these structures and practices intended to support coaches' learning, coaches still reported doing a lot of learning about how to do the specifics of their work on their through experience. Explained one academic dean:

> There is not […] a vibe in the school [that the principal, other academic dean, and I] know everything there is to know about anything. We just basically have the time built in our schedules to actually go find what you need and then to learn about it and then to show it to you. And then what happens is once you show it to one person, I can now teach lots of teachers about questioning because I […] help[ed] one teacher with it. (Interview21)

This independent learning was still supported to some extent by AF because it took place with some guidance and oversight from the coach's coach, because there is time set aside for this

work since it is seen as part of the coaching role; and because the coaches were selected in part for their skills in this area.

Data analysis and planning. The powerful reports available for teachers through the Athena platform, the Data-Driven Plan (DDP) protocol and guide, and the support teachers got through new staff training and optional network-wide workshops about data analysis were important tools AF provided its teachers and school leaders. However, AF leaders believed teachers needed much more support to use Athena in more sophisticated ways because their data analyses could have real instructional consequences. Shelley Thomas, Product Manager of Athena, shared an example of what this could mean for teaching and student learning. She suggested that if a teacher were only to look at reports from Athena showing students' scores by standard, such as understanding a text's main idea, and found the scores to be lower than anticipated, then they might engage in a simple reteach of main idea. However, combining that information with another report that displays the proportion of students answering each question on main idea correctly, "You might actually see, well they're really knocking it out of the park when it comes to multiple choice, but there's something about open ended [questions about main idea] that they're really struggling with," she explained. By clicking on the embedded link to examine the specific open-ended question(s) about main idea on that IA, as well as the scanned student responses, the teacher would have more information with which to diagnose students' difficulties and might discover that rather than main idea, there was another important skill that students needed support with, like organizing a response to a short essay question.

What is more, teachers may need support in learning to combine what can be seen through Athena with other professional knowledge about the academic content and their students. CIO Ferrell observed, "Athena can be illustrative, it can hold up some flags […] but it's

not the whole story. So when you look at the item level analysis and you see the students who chose which distractors, that can start to tell you something about their misunderstanding but you also need to bring in some other context." To make sense of how to plan for instruction in the future, for example, the teacher might need to ask questions like, "Was this a non-fiction passage or a fiction passage, […] how close to reading at grade level are [the students who missed it]?"

Although AFNS provided teachers with access to Athena and the corresponding tools to help support strong data use, much of the guidance in how to engage in careful data analysis and planning took place at their school site in an attempt to support them in learning to make good use of their results in their specific school context. AF's model for building capacity for strong data analysis and planning is therefore intentionally school-based, with the expectation that "the school leaders are taking ownership of developing their Athena capacity for their teachers," Ferrell explained. He reported this approach had met with mixed results in part because some principals were "by nature more data savvy and more data hungry," and in part because some did not realize how much support some teachers needed in learning to use Athena well. Yet despite this unevenness, AFNS leaders believed it was their role to support school leaders so they could support their teachers in this work. Ferrell hypothesized that without building this capacity in the school's leadership, efforts to train teachers in data analysis were unlikely to create school-level capability that would persist. As a result his team planned to provide more training in data use for teachers, but especially to bolster the supports they were providing to the principals, which already included examples of how to structure data days, strong data day agendas, successful team sessions, and training for deans and grade level chairs.

School-based data days at the beginning of each IA cycle were critical opportunities for teachers to learn to conduct useful data analysis and planning. While the format of these data

days varied by school, generally the staff convened to review overall school, grade level or subject-area progress, strengths, and weaknesses relative to their achievement targets and to their peers across the network. At many schools on data day, grade level leaders reported out to the entire school about the results, which, a principal explained, "shows that they're doing a lot of work in organizing all that information [and] that they really know all of their work well. But I think in a way it is also a public act of [saying], 'I'm the one who is ultimately responsible for the performance of all 60 kids in this grade'" (Interview44).

Then, in grade level or subject area teams usually guided by a grade level or content chair, teachers worked to make sense of the IA data (and any other relevant assessment or observational data) and then engage in instructional planning.[57] Working in in teams on data days also allowed teachers to think strategically about how to share their resources and those of the school for the benefit of the students across the grade level. For example, attending two elementary schools' fourth data days, I observed grade and content teams in both schools working collectively to interpret data and to plan for instruction, sometimes drawing on previous years' plans or pulling units or lessons from the shared AF server to inform their work. I also saw grade level teams make decisions together regarding regrouping students for instruction across classes; changing the grade level schedule in ways that they believed would benefit student learning; requesting that principals, principals-in-residence, academic deans, or deans of students take on some of the re-teaching or intervention work with small groups during the next six weeks to enable them to create more refined groups; contacting AFNS' curriculum team to

---

[57] Depending on the needs of the staff, principals or academic deans may also engage their faculty in whole-school data analysis and planning, or support their work by participating in these team analysis and planning meetings in which team members work together to make sense of common misunderstandings and how to plan accordingly to address students' continuing needs. AF middle and high school data days often involve more cross-school collaboration among teachers, or more work with school or network leaders, particularly when schools are new and may just have a single grade, and therefore, single teachers in some subject areas.

clarify expectations about the scope and sequence and next IAs, or to ask for support in teaching

a particular set of objectives; discussing individual students' progress and needs; and sharing

ideas about effective teaching moves. Although teachers and school leaders reported that some of

this work took place at weekly or biweekly grade level team meetings throughout the year, data

days were particularly focused time for this collective work.

Developing teams capable of working together in these ways was something that some

school leaders reported having cultivated intentionally. An academic dean in one of the schools I

observed with particularly strong working teams reported:

> [Building strong grade level teams that work well together and share
> responsibility for students in the grade] has been something that we've been doing
> since the very first year at our school, so that kind of just becomes part of the
> culture, of […] we're working toward a grade-wide goal. So we want to see 100%
> of kids score over 85% on the math test. If that's what we want, then we're going
> to have to dig in together to get there, and help each other to get there. So that's
> where the buy-in comes. I'm not setting a goal for each class, I'm setting a goal for
> the grade and we're all working towards that.

Still, she noted that this shared sense of responsibility did not develop without any tension, in

part because of the school leaders' commitments to both teacher and student growth.

> One question that has come up alongside [collective responsibility for students
> and grade level performance] is whose responsibility is it when kids struggle? For
> example, if kids in a certain class are struggling, who should take the
> responsibility? Should the teacher next door have to take those kids who are
> struggling because that teacher hasn't excelled with them, or should that teacher
> get better? I think that goes back to the short-term versus long-term fixes. We
> always want to do what's going to be the long-term [fix], whenever possible. At
> the same time, we're dealing with kids who have to excel and have to get better
> and these are their lives on the line, so we sometimes have to go towards that
> short-term, let's have that teacher next door help out because otherwise these kids
> are not going to get the education they need. (Interview30)

This example of a network-wide sense of urgency about each child's learning that this dean

reports as guiding her school leadership's decisions in the face of tensions like this—an example

of one of AF's core values, "first things first," which refers to "the needs of our students" (AF,

157

n.d.)—was echoed repeatedly in interviews throughout the network. When making decisions or contemplating changes in practice, again and again people reported asking themselves and others some version of the question, "How might this influence student achievement?"

Formal talent management.[58] Formal talent management practices provide school and network leaders with another set of instruments for instructional leadership, learning, and management. One critical tool for formal talent management was the professional growth plan (PGP), a formal, annual assessment completed for each of the network's teachers.[59] In addition to providing consistent, coherent feedback to teachers across AF, PGPs were intended to formally communicate professional expectations; to structure school leaders' training in important leadership skills; to identify strong or weak performers and to make instructional growth and career development plans for them; and to document poor performance that may require dismissal.

The PGP development process was spearheaded during the 2008-2009 school year by Sarah Coon, then the Director of Evaluation and Organization Development, who worked closely with principals on their design and implementation. The PGP assessed teachers across seven broad areas AF believed to be central to their mission and students' academic growth: Student Achievement and Strength of Character; Instructional Excellence; Classroom Culture; Planning and Data Analysis; Student and Family Relationships; Personal Organization and Effectiveness; and Core Values and Responsibilities. Within each of these categories were a variety of

---

[58] AF's formal talent management and compensation practices have changed greatly since the time of this study. For example, the PGP and Talent Review processes described here have largely been replaced by the Teacher Career Pathway (TCP)—a new teacher evaluation and compensation system AF designed to better develop and recognize excellence among teachers who elect to stay in the classroom rather than assuming other leadership roles, and to further align teachers' evaluation and salary structure with their performance. Although the TCP was only being developed and piloted the year of this study, the program's design provides insight into questions AF was grappling with during these interviews. A discussion of this program is provided in Curtis (2011), Doyle and Han (2012) and Rosenberg (2012). Modified PGPs are still conducted for school leaders, AFNS staff, and operations staff, however.
[59] School leaders also went through the PGP process with their coach using a document modified for their role.

indicators aligned with AF's core values and the Essentials, upon which teachers were rated from 1 (Does not meet) to 4 (Mastery). The PGPs also required reflection and comments on areas of exceptional performance and areas for growth within each category. New teachers were directed to focus on their performance relative to PGP indicators marked as "foundational" while returning teachers were asked to reflect on their progress relative to the previous year's goals.

Mid-year, teachers were asked to engage in a self-appraisal guided by the PGP citing evidence—particularly student achievement data—for their ratings when possible. Next, each teacher's coach or principal added his or her own ratings and comments to the PGP before filling out a more holistic performance summary. The completed PGP was intended to guide a feedback conversation in which they would discuss both parties' ratings and reflections before updating the performance summary together to focus on key strengths, growth areas, goals for learning and development, and the supports necessary for achieving those goals. Principals, coaches, and teachers were expected to return to the reflections and goals in the PGPs to guide their future work together.[60]

After the first year of using the PGPs, AFNS collected feedback about how valuable teachers found the PGP for their own learning and development. Coon submitted:

> [I]n one school, 19 percent said it was extremely or very valuable; in another school 90 percent said it was extremely or very valuable. And that was such a big a-ha moment for me; [I realized] that difference has absolutely nothing to do with this piece of paper that we emailed to them and it has everything to do with the relationship between the coach and the teacher, so that is really where we focused this year.

As a result of this feedback, AFNS made only small modifications to the PGP document itself in 2009-2010. However, they changed the order in which PGPs were completed, with school

---

[60] While PGPs were also collected by the network staff, during the 2009-2010 school year, PGPs were completed in Microsoft Word, which meant the data they contained was not in a format conducive for large-scale analysis. Coon reported that getting the PGP "online" to facilitate analyses of the information collected was a priority for the network in subsequent years.

principals, then deans, then teachers going through the process, instead of the reverse as it had been the first year, so that principals and coaches could experience the process for themselves before leading their teachers through it. AFNS also asked teachers' coaches to allow no more than two weeks to elapse between teachers' self-appraisals and the feedback conversation so that the process was more timely and relevant. Finally, AFNS bolstered the training and toolkit that they provided to their school leaders to support this process. School leaders were now trained in providing strong written and verbal feedback and guidance in addition to a range of supporting tools—including examples of completed PGPs of varying quality for stronger and weaker teachers, videos of more and less successful feedback conversations, and scheduling support.

Because of the intense collaboration among and between teachers and school leaders, not to mention IA results, it would be unusual that a mid-year PGP would be the first time a teacher would be identified as struggling. Yet when coaching and the typical PGP process was inadequate for addressing the needs of a struggling teacher, AF also had a formal Performance Improvement Plan (PIP) process that was intended to clearly identify and communicate a teacher's areas of weakness with professionalism or instructional performance; to ensure that the teacher understood that these issues endangered their job; and to establish a clear 4-8 week action plan that included weekly progress support and check-ins. If teachers still did not improve or if the performance problems were egregious, teachers could be terminated before the end of the school year. Roughly 5 percent of AF's teachers are not renewed annually (Curtis, 2011).

School and network leaders tended to report believing in the importance of transparency in this process for the benefit of morale in an at-will employment environment. Below are two comments typical of these leaders on this topic:

> Our teachers are at-will, so they get offer letters every year. If somebody […] is a potential non-renewal, we need it to be really clear in their PGP how they're

performing and what they need to do in order to improve and in order to stay. [To not do so] is not fair to those teachers and it's not good for the organization as a whole. (Coon)

We don't want the school culture to be, I'm frightened for my job, […] I can get fired at any time. It's already there because it's in the at-will contract. So we want people to feel like, yes, you need to keep improving and be great, but we're not going to just fire you because we don't like you, it's going to be because we together, you and the admin team, have decided this is not the right place. No one has ever been surprised. (Interview21)

For these and other reasons, school leaders almost uniformly reported that they were willing to invest a great deal in their struggling teachers if they demonstrated a strong desire to improve and believed in AF's mission. For example, one principal reported that she and the other leaders at her school tended to distinguish between "will or skill" problems. She found that typically, the PIP process worked well for skill problems. Problems with will, on the other hand, could be more difficult to address, but she reasoned, "Achievement First is not the right organization for all people who are in teaching," and that it was important to have conversations with that teacher about whether or not AF was the "right fit" (Interview19).

Similarly, another principal explained that the teachers that she had dismissed had mainly demonstrated a lack of fit with AF's values. She reported developing strong concerns about their potential as a teacher in her school "basically if they believe anything other than teacher actions can entirely determine student actions and with the right teacher actions the students will act appropriately academically and behaviorally." She also worried if that person did not invest in their own improvement: "Unsuccessful attempts are okay. But if you're getting feedback twice a week and you're not attempting to incorporate it, and you're getting coaching about how to incorporate it, someone demo'd in your room for a week, […] then I'm really concerned." In fact, she reported that she has "only once needed to exit someone for pure failure to improve. Other times it has been mind-sets completely." This principal preferred to dismiss a teacher mid-year if

things were not going well, even though she and her leadership team had to take over in the classroom until a new teacher could be found: "It's difficult to re-staff [mid-year]," she conceded, "but it's better to have no teacher than a teacher who doesn't believe in your kids" (Interview18).

In addition to the perspective on "talent" provided by the PGP and PIP processes, AFNS helped principals to undertake a "Talent Review" in which regional superintendents work with the principals in their portfolio to holistically and systematically assess each of their teachers, deans, and other operations or support staff using an AFNS created Talent Review Worksheet. After sending this review to AFNS at the end of October, AF's Director of Leadership Development Sara Keenan created "Talent Snapshots" and workshops for the principal cohort. The goal is that these snapshots will help principals and regional superintendents to get an overview of teachers who need support or might be developed for leadership; to develop strategy about recruitment or teacher reassignment for the following year; to identify who they want to work particularly hard to retain; and to collect information that helps them to assess the previous year's recruitment and selection process. These talent snapshots were also collected by AFNS, who used the information in similar ways from a network perspective.

The PGP and Talent Review process appeared to offer useful ways to identify and recognize strong teachers who might be groomed for leadership roles within the school and broader network. AF hoped that developing an internal career path will benefit the network in several ways. First, an internal career path could serve as a way to recognize and retain AF's most talented teachers and school leaders. Keenan reported that AF aspired to "paint a career path for talented people" and to make sure that "those most talented people know how much they contribute to their school." The network ran a leadership fellows program that was intended to serve in both of these capacities by identifying and developing promising future leaders—

162

including future teacher coaches, grade level or content area chairs, deans, or principals. These roles and responsibilities came with extra salary beyond the standard salary scale at AF (roughly 10-15 percent above the salaries provided by the surrounding school district to compensate teachers for the extra time they spend working in AF's model), allowing for additional recognition of strong teachers and team members and the extra work that these positions entail. Other strong teachers were tapped to work as leaders within AFNS. Keenan suggested also worked to recognize strong teachers in less formal ways by providing them with more time and attention. This included strategies like spending time observing in their classroom and giving them feedback, working to build a positive relationship by inviting them for a drink or for dinner to discuss future plans, or even just writing a letter, card or email to acknowledge their work. AF also worked to connect these teachers to external professional or leadership development opportunities and to fund these ventures, or encouraged them to share their talents through AF-wide PD day or by becoming professors at Teacher U (now Relay Graduate School). Finally, this sometimes involved having explicit discussions about how to support a teacher's life goals in addition to professional ones.

In addition to recognizing strong teachers, these talent management processes and leadership development programs were aimed at providing AF with a steady stream of leaders who were already successful with, and knowledgeable about, the instruments and practices described earlier in this essay as turnover took its toll and the network continues to expand. Indeed, by the 2009-2010 school year, the vast majority people hired for school leadership positions were internal candidates. For example, in the year following this study, due to turnover and expansion, AF anticipated having 5 new principals, and all 5 had previously been successful

academic deans, and before that, effective teachers. The previous year, the network had hired

approximately 20 deans, and all but 1 was an internal hire.

POSSIBLE CHALLENGES TO AF'S CENTRAL GOALS

AF's infrastructure, and the ways it was developing and used, was still nascent and uneven in the

spring of 2010. In interviews, some network leaders, school leaders, and teachers expressed

frustration with the quality or execution of some pieces of the infrastructure and its repercussions

for their instruction and student learning. Others found that managing the social dynamics

involved could be difficult because of the still-developing professional skills of the often young

and inexperienced people working in AF. However, I believe that AF faces two central

challenges as they work toward their fundamental goals of supporting poor and minority student

success at the scale of an urban district to inspire broader educational reform. The first of these is

related to the ways in which success in teaching and learning has been defined in the educational

contexts in which AF is situated and in AF itself, and the second to the sustainability and

potential for replication of AF's model. Below, I briefly sketch some of the primary challenges

in each category and elaborate upon this discussion in a companion essay (Rosenberg, 2012).

First key challenge: Definition of success in teaching and learning in AF and beyond
While the increased accountability to which AF's schools agree to secure their charters provides

them with crucial flexibility that allows them critical space to develop an infrastructure of

practice and the supports necessary for it to be used well in practice, this heightened

accountability may also be a liability. Student performance on standardized assessments in the

states in which AF operates are critical metrics of the network's success for both the charters'

authorizing agencies and important funders, and as such have come to deeply influence the heart of AF's own definitions of teaching and learning quality. The AF scope and sequence and IAs were largely derived from these standards and assessments, particularly through 8th grade. The Cycle and the Essentials represented a vision of teaching that AF's leaders have found supported relatively new teachers in helping their students to perform well on these IAs and end-of-year assessments. Many other key aspects of AF's infrastructure were devoted to ensuring teachers developed competence in supporting, or were evaluated with respect to, students' performance on these exams. With the states' assessments, content standards, and proficiency standards providing the central organizing constructs for the internal and external assessment of the network's accomplishments, AF's ability to support their students in meeting with success in higher education and beyond depends on the extent to which these external tools serve as reasonable proxies for students' progress toward these goals—or on the organization's ability to help teachers to learn to teach in ways that help students to engage in learning more ambitious material in addition to performing well on the state assessments.

In the United States, state standards and assessments have not been known for their strengths, even in the relatively affluent states of New York and Connecticut. In a report for the Fordham Institute, Schmidt, Houang, & Shakrani (2009) argue that U.S. standards are neither rigorous nor coherent, largely because of their lack of "focus" or the concentration "of instruction at each grade level around a reasonably small number of topics," relative to high performing nations (p. 24).[61] Instead, the standards tend to cover a plethora of topics in each grade so "each topic is covered superficially—and, often, repeated grade after grade" (p. 24). Rather than organizing topics "within and across grades [in a manner] consistent with the

---

[61] In previous work, Schmidt and Houang (2007) established that cross-national mathematics achievement was related to the coherence, focus, and rigor of a nation's standards.

inherent logic of the [academic] discipline," state standards in the U.S. have instead included

"topics put together arbitrarily in a process governed more by politics than substance" (pp. 24-

25). These numerous, discrete and poorly organized standards tend not to be carefully aligned

with the standardized tests that are supposed to assess them, and the tests themselves are

frequently criticized for emphasizing low-level skills. States also establish the scores required for

students to be deemed "proficient" in particular standards or groups of standards and these

proficiency targets may also be set quite low (de Mello, 2011).

Within AF, some leaders were not deeply concerned about the threats of preparing

students too narrowly given AF's strong focus on student achievement as measured by interim

assessments and state standardized tests. For example, CIO Ferrell pointed to the fact that AF

designed their own IAs and distributed them at the beginning of each IA cycle, so even though

the IAs were built to prepare students for the state exams, AF was able to exercise significant

control over they ways that teachers interpreted the standards and the kinds of tasks that students

needed to be able to complete to demonstrate understanding of the standards to perform well on

the IAs. An elementary school principal I interviewed expressed his view that AF's strong focus

on achievement was not troubling because of the organization's concomitant commitments. He

first explained that he believed that the assessments themselves were "excellent indicators" of

students' progress toward the network's goals for them, because "almost all other indicators are

reflected in achievement results, in that there is such a high correlation between all the other

pieces of data that we collect and achievement results that if we're going to settle on one number

to look at, achievement results still make the most sense." Nonetheless, he continued,

> We also care very, very deeply about building a school culture that is really warm
> and safe and loving and supportive and has incredibly high expectations for
> student behavior, and a school culture in which kids always understand why we're
> doing what we do. I hope that any kid, in whatever is the appropriate language of

> how old they are, could explain how stuff we do at [our school] is related to climbing the mountain to college. […] There are so many elements of school culture, there's been a tremendous amount of time dealing with it, but I think that it is a sort of artificial distinction to say that is a totally separate program that has separate ways to measure success.

Still, he noted, "I do think there are some other school culture projects that need to take on a little more of a life of their own that aren't just about supporting the academic part of it." For example, across the network, he reported people were working to figure out how to improve in their "teaching and modeling and helping the kids understand character growth in a deeper way than as just earning rewards or complying with directions, [but also] what it means to be a citizen, [or] to be somebody where they deeply believe in giving back to their community for their whole life." He suggested that the way that AF used test scores also helped to make them less concerning to him: "In our mind it's not the end-of-year scoreboard of how we did, but this constant real-time data that helps us adjust how we're doing" (Interview44).

Yet I also encountered a significant minority of network leaders, school leaders, and teachers who were more skeptical about the rigor of the state standards, assessments, and proficiency targets and therefore the extent to which students' scores on standardized assessments reflected meaningful progress toward success in higher education and beyond. Some of these individuals shared concerns about the ways these features of their environment shaped the quality of instruction via the scope and sequence and IAs, particularly in an organization employing so many novice teachers and instructional leaders, and particularly in mathematics.

Ironically, though AF's elementary school mathematics scores were stronger than those in literacy, the school leaders and teachers I spoke with tended to highlight their concerns about the influence of the scope and sequence and IAs on the quality of the network's elementary math instruction much more frequently than they did in literacy. This may be partly because the

167

network had adopted the F&P Continuum in all of its elementary schools, which, as it was being used in AF, seemed to provide teachers and instructional leaders with a reasonably strong instructional system to guide, integrate, and organize important aspects of their work in reading instruction, as did the UEI's STEP assessment system used in some elementary schools. In addition, with about three hours daily for literacy instruction and an additional writing block (usually 45-60 minutes long), teachers often reported feeling that they had the time to engage in the many different kinds of instruction they believed their students needed in addition to preparing them for IAs and state tests—including specific times of the day set aside for various pieces of a "balanced" literacy approach that varied somewhat by grade, school, and region but including things like additional phonics and other "reading skills" work; textual analysis (basic reading comprehension skill development that many teachers and school leaders referred to as test prep); "literature" time; independent reading time; reader's or writer's workshop, or direct vocabulary or grammar instruction.[62] They also reported that these aspects of literacy instruction were relatively easy to align with the network's scope and sequence.[63]

Mathematics, however, did not receive as much instructional time. Generally, math lessons were about 60 minutes daily in the elementary schools, though some teachers worried that some of these minutes were lost as students were regrouped for instruction. Students also participated in some math instruction during a separate, daily math meeting (usually 20 minutes)

---

[62] Different schools drew on specific programs to structure parts of this work. For example, in addition to the programs already mentioned, some schools used Direct Instruction or Reading Mastery for students who still needed support with decoding and fluency, Wordly Wise for vocabulary development, or Junior Great Books or a version of "Reading Mania" developed by a KIPP teacher to guide literature instruction. Textual analysis materials were generally provided by AFNS. Many teachers and instructional leaders reported that while they followed some of these programs closely (like F&P, or some of the basic fluency and decoding programs), others simply served as references for them as they developed their lessons. Several school leaders and teachers I spoke with were also actively involved with the Lesley Literacy Collaborative or with the Teachers College Reading and Project, and these perspectives also featured into the network's literacy work, though unevenly.

[63] An academic dean noted that with textual analysis, "it's basically just understanding the text and answering questions, [so] once you know what main idea is, once you know what fact and opinion is […] you can get better [at it], and those IAs don't change, they're basically the same skills across the year."

that was generally focused on cumulative review or distinct skills (e.g., work with the calendar, time or shapes for young children); and a smaller subset participated in math intervention programs. However, with no instructional system comparable to the F&P Continuum or STEP for math AF lacked a framework and sets of tools that might support relatively new teachers in making connections among the numerous discrete standards included in the state standards IAs.

This state of affairs influenced the curricular choices of many elementary schools, with many selecting the enVision mathematics curriculum as a reference for their teachers—not because they thought it was necessarily the best program, but because teachers could pull units from enVisions out of sequence to fit with AF's scope and sequence. Other schools crafted different solutions. For example, an academic dean explained that she and others on her staff were committed to the ideas about mathematics instruction embedded in the Investigations curriculum because of its emphasis on developing deep understanding of the content. However,

> Last year in second grade, we wanted to continue with the Investigations program but it didn't really match with what the IAs were, so we were getting this really solid math teaching for three weeks and then really broken off test taking skills before the IA. So what we did this year to combat that is to say, all right, well, we're not going to take the Investigations into it because it's a program […]. [Instead] we're going to figure out how we can take the IA and turn it into a thinking process of a test […] So instead of teaching kids double digit addition, you know, put the number at the bottom, carry the one, […] they learn how to do it through place value and through manipulatives and things that you would do in an Investigations-style lesson with […] them being able to do it on paper.

She reported that this was frustrating, but that the predominance of the state standards and assessments was undeniable: "It is the reality of our education system right now, and so we can't move beyond that. At the end of the day, this is how they rank your schools. We're not even on the map until our [Connecticut Mastery Test results] come out" (Interview21).

Regardless of the math texts selected by the school, several school leaders and teachers reported that the larger problem was that as Schmidt et al. (2009) suggested, the state standards

and assessments, and AF's corresponding instructional tools, pushed teachers in the direction of

"coverage" in mathematics rather than the cultivation of deeper understandings, even when

teachers and school leaders were aware of and uncomfortable with this. For example, an

academic dean explained,

> [Y]ou have six weeks, and you have these three blocks of standards that you need to cover. You need to do fractions, you need to do money, and you need to do graphing. Fractions could take that whole six weeks, depending on where the understanding is, but knowing that you kind of have to move through it at that pace… So I think that's someplace where we sometimes shortchange on the understanding because of the pace that we have to keep up. [...] [Y]ou know that you're going to have three questions on money and if you never got to the money unit […] you're not sure how your kids are going to do. (Interview30)

An AF teacher at another school who had been teaching for nearly a decade expressed similar

concerns. She clarified that she was "no expert" on mathematics teaching, but observed:

> [T]he enVision [curriculum] and the scope and sequence both encompass so many different topics of math, which I think are extremely important. But especially for second grade, I'm wondering if it would be more beneficial to the students and just push them further in their level of mastery if we didn't so much even bring up [topics like] time, maybe, or bring up geometric solids or graphing or something like that, where we spent the entire year on like two main things, maybe numeration and number sense, and addition/subtraction, that would be it.
>
> Because we're finding in third grade, I'm watching them count objects to add them together; where there's a set of five things, they should be able to look at that and know it's five, and they're not able to do that; they're not able to see ten objects, and if they count the ten and then there's a line directly underneath that's—they're lined up perfectly, they have to then recount the second line. And I think all of those should be mastered and solid by the end of second grade and they're not. And I don't know if it's because we are trying to teach so much at one time we're not giving them a chance for mastery.
>
> I was [looking] at the third grade scope and sequence for IA five and there were a lot of new things and one of the biggest things is fractions, and it's equivalent fractions and ordering fractions, and to me that's really difficult, really difficult. So I would like to spend even more time than just the six weeks just on fractions, but I wasn't able to. […] So, again, I'm not an expert, but that's just what I'm seeing. (Interview43)

This teacher's comments usefully highlight the difference in the approach the network was developing for adult learning through their "coaching to mastery" strategy through which they focused supporting teachers in developing on one or two major strains of their work at a time, and the network's approach to elementary mathematics instruction.[64]

At the same time, even educators and leaders with concerns about the state standards, assessments, and the AF instructional tools built from them appeared to take them seriously for several reasons. First, even if the state tests set "a relatively low bar," regional superintendent Marc Michaelson pointed out that this bar can be difficult to meet with consistency. He admitted, "I see us still having a ways to go to getting even to that base level of performance in some areas." In addition, people throughout the network believed it was necessary to have some standardized tools with which to assess their progress. The comments of an academic dean reflect the ambivalence about the tests expressed by many of her colleagues:

---

[64] These struggles with elementary mathematics absent strong instructional systems in mathematics—like the F&P Continuum in literacy—suggest it might be useful to draw upon or collaborate with others who are deeply engaged with similar problems. For example, Professor of Teacher Education and Senior Advisor to the Boston Teacher Residency Magdalene Lampert is working with several colleagues to study and develop innovative models for supporting novices in engaging in "ambitious instruction." Ambitious instruction is that which is deeply tied to content and "deliberately aims to get all kinds of students […] not only to acquire, but also to understand and use knowledge, and to use it to solve authentic problems" (Lampert & Graziani, 2009, p. 492). For example, Lampert, Beasley, Ghousseini, Kazemi, and Franke (2010) describe their work in several university teacher education settings to develop instructional tools and practices to support novice teachers in learning ambitious mathematics teaching— including the development of instructional activities or "'chunks' of teaching that maintain the complexities of practice while simultaneously providing manageable, structured routines that constrain instructional choice" (p. 135). These instructional activities "are intended to maintain complexity in that their structure encompasses an instructional sequence that enables a teacher to address a particular instructional purpose (albeit at a range of different levels) in principled, ambitious ways. The predetermined, stable structures of the instructional activities we are using constrain the set of decisions a beginning teacher (or experienced teacher, for that matter) must make during their enactment. [The four instructional activities they have developed] all target teaching and learning in the domain of number and operations at the heart of elementary mathematics and can be used to accomplish multiple learning goals in lessons across the elementary spectrum. Our hypothesis is that this set of activities will serve as a productive starting place for novice teachers, enabling them to develop broadly applicable skills and knowledge "(p. 136). These instructional activities will "structure what teachers and students do together to bring about an intended learning goal," but at the same time will "leave room for teachers to create teaching in response to what is displayed (p. 137). Ultimately, the authors hope to contribute to the construction of "building a theoretically and empirically grounded instructional system for elementary mathematics" (p. 136) that might support teachers in and beyond their first years of teaching, and teacher educators (defined broadly) in supporting them with this work.

> The bar is so low on the state test it's hard to know what it is really preparing kids for. I think that that's just a challenge of the state test. You can get a 3 [out of 4, which designates proficiency in New York] on the math test and really that doesn't mean that you know a lot of math, it means basically you figured out how to do those types of problems, and you're able to do that, especially because so much of it is multiple choice. I think the reading test is maybe less of that but it doesn't tell you that you're a really strong reader as much as [that] you're able to answer these types of question about the text.

Nonetheless, she also emphasized a similarly prevalent view that having a standardized way to measure progress and compare results is essential for the success of the network, the school, and for individual teachers:

> I definitely think that it's better than nothing. I was here before we had IAs in first grade, and I was like, I just want some standardized way of knowing that we're progressing. I think that that's really the strength of it, there's a standardized way of knowing across the schools that we're moving towards a goal. I think that learning and education is really [complex], so to say that one test or a battery of tests can show you if kids are prepared for college is probably a little bit misleading. But to me it's better than each teacher creates a different test and we actually have no idea if what you're testing and what you value is the same as what I am testing and what I value. It's also a formalized way to keep the rigor high across the network. […] I think it doesn't capture everything that kids need to be successful at college, which is basically the broad goal we're working towards.
>
> Interviewer: And as you work with teachers, given that the state test and the IAs are helping to direct your work, do you ever encounter tensions between your work toward those broader goals and work that is measured in the interim ways?
>
> I think definitely, I think that's just something that we're facing across the country. Are these tests the best measurement of what our kids know? Or what our kids can do, or what we really want kids to be able to do as they go to college? But I've actually heard teachers [who have] said, I really didn't believe in this before, but I'm finally seeing how you can use assessment to really guide what you do as your next steps. […] Because I think [being a teacher is] a very nebulous position in some ways, and different people measure [success in teaching] in different ways. So I think it's very rewarding in that sense that teachers can see: I saw this, I analyzed the data, I made a change, and I'm seeing different results. And I think that that's a very positive part of what we've created here. (Interview30)

In these comments, this dean pointed to ways that these IAs and state assessments, while imperfect, helped AF to manage some important challenges that have typically plagued school

systems in the United States. The IAs and state assessments, together with other resources that are part of the infrastructure AF has built, helped to limit some of the uncertainty of teaching by firmly establishing the network's priorities, measuring their progress toward them, and using the results to adjust their plans and motivate further work.

The pervasive awareness about some of the affordances and problems associated with building an infrastructure evaluated largely by state-defined standards and assessments—together with a serious commitment to student success—meant that many AF leaders and educators were actively involved in attempting to manage some of the risks that this might pose. For example, this concern is one of the reasons that AF's network-wide balanced scorecard, the school-level AF Report Card, and the new Teacher Career Pathway system all emphasized a range of academic and other outcomes, and that coaches were deployed to help teachers to focus on the bigger picture of student development as well as achievement. Many interviewed AF personnel felt that in part because of AF's extended day and school year, they were able to strike a reasonable balance between the demands of preparing students for standardized assessments and for the more robust goals they held for them. Others pointed to the strength of developing an AF K-12 pipeline, meaning that they could begin building knowledge and skills not required until later grades in earlier ones, even when not called for explicitly by state standards or exams.

Second key challenge: Sustainability and replication of AF model

Earlier in this essay, I reviewed Wilson's (2008) forceful argument that the small number of high-performing, No Excuses charter schools in Boston he investigated relied heavily upon the Herculean efforts of an elite group of teachers for their success. He contended that the adoption a "powerful instructional system" (p. 36) would allow schools and districts to support large

numbers of teachers from a wide variety of backgrounds to be successful with their students and to avoid the enormous challenges of scale, sustainability, and replication that such a "rare human capital" model entailed. These questions about quality at scale are essential ones for AF because of their commitment to supporting their low-income and minority students to achieve at high levels and their desire to prove that it is possible to do so at the scale of an urban district to inspire broader reform. These goals require that AF's model be both internally sustainable, meaning that its leaders are able to maintain educational quality as they grow, and at least modestly replicable, meaning that their approach is one that may be adopted in whole or in part by other charter or traditional educational organizations.

The AF case suggests that in their quest to set a minimum standard for teaching and learning across classrooms and schools and to support their practitioners in growing beyond this base level, the network adopted strategies that relied heavily upon both a thoughtfully designed infrastructure of practice that goes far beyond that proposed by Wilson *and* a staff with unusual strengths. At least within AF, the development, refinement, and productive use of the infrastructure of practice they were building appeared to enable them to support large numbers of relatively new teachers in supporting their students in achieving at levels at or above their state's suburban peers, but also required them to continue to attract novices, early career teachers, and instructional leaders with exceptional skill or personal resources.

The reciprocal and interdependent relationship between the individuals who AF hires, developed, and retained; the systems built to supervise and support these people; and the professional relationships and organizational culture that helped these various strategies to cohere therefore leaves AF vulnerable to many of the "rare human capital" concerns that Wilson raised in his essay despite their rapid development of a remarkable infrastructure of practice.

AF's aggressive growth as they add grade levels to their schools and schools to the network, the

teacher and leader turnover typical of large urban school districts,[65] and the movement of

successful teachers into key leadership positions early in their career means that AF has to hire

significant numbers of teachers each year. The extent to which AF is able to support reliably

strong teaching across the organization will depend upon whether or not AF is able to either (a)

retain or find new teachers with the competencies AF prizes, and to support these novice and

early career teachers and instructional leaders with sustained learning opportunities for their

continued growth, or (b) elaborate or improve their infrastructure in such a way that it is less

dependent upon the strength of these people.

With respect to retention, the question of the sustainability of the work of teaching or

school leadership within AF is an important one that will receive more attention in this essay's

companion. In brief, however, I found that AF's infrastructure, and the opportunities for growth,

collaboration, recognition, the exercise of discretion, and participation in organizational

improvement that it afforded—not to mention seeing evidence of student learning—helped to

support some teachers and leaders to experience their work at AF as inspiring and sustaining

instead of draining despite what can often be difficult work and long hours. But others were

unsure about how long they would be able to continue to work in an environment that demanded

so much of their time, particularly as they began to think about starting families. And still others

began to feel that as they gained experience as teachers or leaders, the infrastructure that AF built

to support novices in learning to do their work came to feel constraining in addition to

supportive. If such individuals opt to leave AF en masse, it may be difficult for AF to avoid a

cycle in which their dedication to supporting novices makes them more likely to loose precisely

---

[65] Sarah Coon reported that of AF's peer networks, AF's teacher retention rate is the highest. AF retains just over 80 percent of their teachers each year, and the retention rates are in the high 80s for teachers who are asked back.

the slightly more experienced people who may be best positioned to call for and contribute to the elaboration of their infrastructure—requiring them to hire and provide intensive support for even more novices and beginning the cycle again.

AF's focus on continuous improvement and their core mission suggest that over time, AF will elaborate and improve their infrastructure so it supports improved teaching and learning across the network and perhaps enables a greater range of teachers—particularly those with less time to devote—to accomplish a great deal with students working within the systems they build. Yet the unique characteristics of the people that AF employs and the importance of relational and interpersonal work in AF for supporting teachers in using the infrastructure's resources to achieve a minimum standard of instruction and beyond raises serious questions about how much even improved systems might accomplish without them. The importance of the role of the coach in the network helps to highlight some of the limitations of strong systems without also having leaders to help practitioners to learn to use them in the particular contexts of their work. For example, coaches are meant to help a teacher, academic dean, or principal diagnose his or her areas of weakness as a practitioner, and to help them to diagnose the source of their students' or teachers' struggles with specific areas of their work. The coach can provide direct assistance or help to connect the practitioner to other resources within or outside of the network or help the practitioner to address other parts of their practice that may be getting in their charges' success.

In discussing the leadership fellows program for teachers, Freidus, Director of Teacher Leadership, emphasized the importance relational work: "A lot of the program has been built on relationships, I'm really comfortable with that. I think relationships are the heart of educational work and it's really hard to develop people or learners of any age without relationships." However, as she contemplated the fact that the program was slated to nearly double in size the

following year from 23 to 43 participants in order to reflect network growth and leadership needs, she noted, "this [issue of relationships] becomes a question as you scale. How do you make those relationships meaningful and sustainable at the same time? I think that's one big challenge. I think that's the largest challenge, honestly." She noted that the challenges of relational work and scale also arose in the relationships between AFNS and the schools. "We at network support have a lot of resources and are really, really excited to share them with schools and to provide schools with what they need one-on-one," she explained. However, "we're not [going] to do that kind of site-based work without an invitation from a school because there are questions there around ownership. And people aren't likely to reach out to us for that […] unless we already have those relationships."

This importance of solid teachers, excellent school leaders, and relational work also has repercussions for the utility of AF's model in the broader reform community as well. While AF is dedicated to sharing their work with others by inviting them to visit and observe, answering questions, sharing materials, or permitting researchers and policymakers access to the organization—and through more formalized partnerships, as with the school leadership residency programs they now run in partnership with two of their host districts in Connecticut—the organization's leaders are widely skeptical of attempts to reduce their approach to a bullet-point list to be adopted by others, and of how far their innovations will travel without smart, committed teachers and exceptional school leaders. For example, while McCurry thought that many of the critical elements of AF's infrastructure could be transferrable to a traditional district context, he commented that it would be challenging to do so without taking teachers from the top half to the top third of college graduates. Referencing the McKinsey report, he noted, "In America, we traditionally get our teachers from the bottom half or bottom third of the high

school and college achievement pool, and we need to be getting teacher from the top third, and that's what other countries do." Acknowledging that others might say it was impossible to recruit enough of these teachers to the profession, he noted, "I don't know if I agree with that. Top people don't want to go into teaching because of the conditions, and not necessarily because of the pay." Still, he noted that even with strong teaching candidates, it would take time to build a core of strong instructional leaders to spearhead this work in other contexts. McCurry reported that if asked to immediately start 20 new AF schools, he would reply:

> Great. Find me twenty great school leaders and I'll do it. I think that's the challenge. […] Giving okay people good tools is only going to get okay results. […] I think we need to rapidly build up that tier of teacher coach. I mean, the good news on that front is a strong teacher in their third or fourth year can be a coach of other teachers. But they need to have had coaching of them to get there. I think that's actually something the districts can rally behind too; our decision-making on people is not dependent on the number of years of experience they have, it's how good they are. But it takes two or three years in a good environment with good coaching to be able to be a good coach yourself. That's always the thing I see when people try to scale up a reform across the city, I always take a step back and [say] yeah it might have marginal impact, but they need strong people to make it happen. If I were working with the city [say] well, the first thing I would do is find a group of schools where you can really do this intensely to build your cohort of coaches and slowly, slowly spin that out so that you don't kill off a program because people don't think it works. Which is [why I always tell people], use our curriculum, can you use Athena, can you use all this other stuff, and I worry that people are going to do it and say well it doesn't work, and it doesn't work if you don't do all these things along with it.

This perspective suggests that the pace at which AF is able to scale and disseminate its approach depends on the rapidity with which they can build the expertise of their instructional leaders, with school leaders and other coaches serving a particularly essential role in establishing the professional relationships AF's infrastructure requires.

Finally, AF's strong organizational culture may limit the transferability of elements of AF's model to other organizations as well. Stigler & Thompson's (2009) essay exploring some of the challenges of knowledge creation, accumulation, and use for teaching introduces the

concept of "cultural traditions of teaching," or traditions that are "defined by a shared set of goals, assessments, theories, methods, and contextual considerations, all knit together in relatively stable cultural routines." They argue that it is much more difficult to share practitioners' knowledge across rather than within cultural traditions, "primarily because the cultural theories, beliefs, values, and existing routines are already shared within a teaching tradition" (p. 446). Because effective teaching is highly contextual, they argue, "teacher must have theories they can use to think through teaching situations, figure out which of the many strategies might be most appropriate, and then make a judgment about which one to implement and how" (p. 447). Within a cultural tradition, such adaptations are more likely to be productive since "when someone in the same tradition borrows an innovation developed by a colleague, they are more likely to 'fill in the blanks' in the correct way, fitting the innovation into their own practice in an appropriate way." On the other hand, "when sharing across teaching traditions, innovations are more heavily filtered by existing cultural routines," and are therefore more likely to succumb to fatal local adaptations (p. 446). Strong coaches or other instructional leaders might help to mediate some of these challenges, but as noted above, the need for strong instructional leadership puts limits on the rate at which innovation may travel within and outside of AF.

CONCLUSION

Three questions guided my inquiry, though I focused primarily on the first two and only on data from AF for this essay.

(1) What are the key instruments—or the social, curricular, educational, political, or material tools designed in influence instruction—that high-performing CMOs use to define, develop, and manage instructional quality within and across schools? How are they selected, designed, and used, individually or together?

(2) How do CMOs organize to manage their relationships with the environment? How does the environment appear to interact with the organization's design, culture, and operation to influence the instruments and how they are used?

(3) What role does organizational learning appear to play in developing, sustaining, and institutionalizing the network's efforts in these areas?

I found that AF identified or developed an impressive range of instruments and arrangements to support teaching quality across the network that reflected both individual and systemic strategies for teacher and teaching quality support and management. Individualistic strategies included recruitment, retention, and dismissal policies. Professional development strategies aimed at improving individual teachers' knowledge and skill in teaching also fall within this category. However, AF's efforts to ensure that their teachers reliably met a minimum standard of practice, and then had the resources to improve beyond that also led them to develop increasingly systemic approaches to teaching quality. These included AF's work to define a common picture of instructional excellence and to develop, coordinate, and institutionalize a set of tools, systems, and organizational designs to support teachers in developing skill in this vision of their work, and to continuously assess and improve their practice and students' learning.

Notably, as they were used within the network, these strategies were deeply integrated to support greater reliability in teaching quality within and across classrooms and schools. In fact, each of these strategies or tools supported the development or productive use of the others and formed an infrastructure for practice. The network's tools and practices for teacher recruitment and selection were aimed at identifying and hiring individuals with the mindsets (and time) to do "whatever it takes" to help their students to achieve, including agreeing to work collaboratively on and with AF's common instruments, to reflect on and accept feedback about their practice, and to make subsequent changes to their work. This process yielded a staff that tended to be deeply committed to the organization's mission to support low-income and minority urban

students with developing the academic and character skills they need to be successful in college and beyond, and to do so at the scale of an urban district. These hires were willing to employ the organization's no excuses and instructional approach in pursuing this work, and tended be reflective and receptive to feedback with a strong desire to improve and a willingness to work collaboratively. Interviewees' comments about recruitment, selection, and their results suggest that as a result, staff members were typically motivated to work together to learn to use, improve upon, or create, the tools and practices that might support instructional quality across the network. In these ways, the "individualistic" strategy of hiring appeared to serve as a resource for the more systemic approaches to teaching quality.

AF's instructional tools—including the scope and sequence and IAs, the Cycle and Essentials, and Athena and the associated protocols for analysis and planning—provided increasingly coherent guidance to these relatively new but enthusiastic and committed teachers as they learned to do their work. These tools helped to frame common problems of practice around which teachers could collaborate as well as the approaches they might use to manage them. The way that the network organized its committed staff, with AFNS and the schools having a formal relationship of mutual accountability, and with a rich and overlapping set of leadership positions within and across schools, fostered strong professional relationships characterized by trust across the network. These relationships served as crucibles and catalysts for learning to use (or improve, or add to) these tools in the specific contexts of teachers' work in collaboration with peers and more knowledgeable or experienced others. They also provided leaders with sustained access to their colleagues' work and with ways to monitor, learn from, celebrate, or remediate teacher and instructional leaders' practice, or eventually to fire teachers who do not "fit" or improve.

AF's dominant culture was created by and animated each of these aspects of AF's infrastructure. By organizing much of the professional work that took place in the network so that it happened in close collaboration with like-minded peers or a coach, structuring clear opportunities for professional growth and evaluation, emphasizing transparency in practice and data about practice, and focusing on continuous improvement in the effectiveness of individuals and groups, AF's staff helped to reinforce shared dispositions and norms and to bolster their dominant culture. AF's culture, in turn, helped to enable practice. Shared values and norms helped people to develop the sense of trust, mutual accountability, and the assumption of best intentions that people so often referred to in their discussions of what enabled productive coaching, collaborative work with colleagues, or work across the network and the schools. Indeed, AF's dominant performance, growth, and team-oriented culture seemed to serve as a kind of connective tissue among the other more tangible features of AF's developing infrastructure of practice. In addition to linking different parts of the infrastructure together by helping to focus priorities and norms for action and interaction, the dominant culture within the network helped to shape the ways that individuals received feedback on, made sense of, and responded to features of their work in the infrastructure's interstices or when they need to adapt elements of the infrastructure to fit their particular needs.

It is worth highlighting how vastly different this approach to education reform is from the state-level standards-based reform efforts that are more frequently referred to as "systemic" in education. Standards-based reform is intended to create a coherent system of instructional guidance by holding schools accountable for student performance on standardized tests aligned with state content standards. Cohen and Moffitt (2009) describe these instruments as creating a framework or "exoskeleton" (p. 10). Theoretically, other more interior parts of the educational

system, including teacher education and professional development programs would have incentives to pull themselves into alignment with these frameworks, as would local schools and teachers who would seek out these and other resources to enable them to perform effectively (Smith & O'Day, 1990). Yet as Cohen and Moffit point out, this approach glosses over a critical irony: these reforms require that the same institutions and individuals whose previous problematic performances compelled their reform develop the capability to take on the difficult work of filling the enormous gaps that remain between this exoskeleton and practice. If this is to result in system-wide to reform, this work would have to occur with great consistency across teachers, schools, districts, and states. Yet because schools and districts with high-poverty, high-minority student bodies are also the most likely to employ a staff with low levels of training, support, or experience and a high degree of turnover, they are also the least likely to be able to learn to use these tools effectively and consistently.

In contrast, AF's infrastructure and the ways it was deployed in the network organized professional development, advancement, and learning to develop instructional and leadership capabilities in the talented but inexperienced people they hired. By establishing and engaging educators throughout the network in the pursuit of common goals using shared tools, this infrastructure and organizational culture provided mechanisms for intense discussion, support, and oversight for the core work of schools and classrooms. It also allowed school leaders and teachers to experiment with ways of organizing school and classroom work; to tailor shared resources to the meet specific local needs and address common problems of practice within widely agreed upon boundaries; and to learn together from these efforts.

In all of these ways, AF provides a rich and generative new model for thinking about what organizing for quality teaching and learning might look like in the US context and for what

the roles of both individualistic and systemic strategies might be in this work. However, two of the greatest strengths of AF's infrastructure—including (a) the way it is deeply focused on and organized around supporting student achievement defined in a particular way, and (b) its integration of individualistic and systemic strategies, and the organizational culture that animates them—may also prove to include some of its greatest challenges. Indeed, this conclusion foreshadows a critical theme in the following essay in which I build upon the description of AF's infrastructure offered here to examine the way that it tended to be experienced and used by the practitioners working on and within it. I found that AF's infrastructure functioned as a framework and scaffold for individual and collective learning; as a safeguard that helped to ensure network staff and students reliably realized at least the short-term goals the network held for them; and as a haven for young professionals from some of the dysfunctions of the broader system. Yet for more experienced teachers and leaders, AF's infrastructure also sometimes came to feel like an impediment to their professional growth.

References

Abdulkadiroglu, A., Angrist, J., Cohodes, S., Dynarski, S., Fullerton, J., Kane, T., & Pathak, P. (2009). *Informing the Debate: Comparing Boston's Charter, Pilot and Traditional Schools.* Boston, MA: The Boston Foundation.

Achievement First. (2009). Strategic Growth Plan, April, 2009. Appendix D, Sample Charter Management Agreement. Provided by Achievement First.

Achievement First. (2009b). Strategic Growth Plan, April 2009. Provided by Achievement First.

Achievement First. (2009c). 2009 Teacher Criteria Map. Provided by Achievement First.

Achievement First. (2009d). A Guide to Data Driven Planning with Illustrated Examples. Provided by Achievement First.

Achievement First. (2010). Shared Beliefs and Practices Working Draft (for discussion). Spring, 2010. Provided by Achievement First.

Achievement First. (2010b). AF Coaching the Essentials, January 26, 2010 draft. Provided by Achievement First.

Achievement First. (n.d.). Achievement First Approach: Core Values. Retrieved March, 2010 from http://www.achievementfirst.org/our-approach/core-values

Angrist, J.D., Cohodes, S.R., Dynarski, S.M., Fullerton, J.B., Kane, T.J., Pathak, P., & Walters, C.R. (2011). *Student Achievement in Massachusetts' Charter Schools.* Boston, MA: Center for Education Policy Research, Harvard University.

Auguste, B., Kihn, P., Miller, M. (2010, September). Closing the talent gap: Attracting and retaining top-third graduates to careers in teaching. McKinsey & Company.

Ball, D.L. & Cohen, D.K. (1999). Developing practice, developing practitioners: Towards a practice-based theory of professional education. In G. Sykes and L. Darling-Hammond (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 3-32). San Francisco: Jossey Bass.

Barr, R., & Dreeben, R. (1983). *How schools work.* Chicago, IL: University of Chicago Press.

Berends, M., Bodilly, S.J., & Kirby, S.N. (2002). *Facing the challenges of whole-school reform: New American schools after a decade.* Santa Monica, CA: RAND.

Betts, J.R., & Tang, Y.E. (2008). *Value-Added and experimental studies of the effect of charter schools on student achievement.* Seattle, WA: CRPE.

Bidwell, C., & Karsarda, J. (1980). Conceptualizing and measuring the effects of school and schooling. *American Journal of Education, 89*, 401-431.

Bierly, P.E., & Spender, J.C. (1995). Culture and high reliability organizations: The case of the nuclear submarine. *Journal of Management, 21*, 639–656.

Birman, B., LeFloch, K. C., Klekotka, A., Ludwig, M.,Taylor, J., Walters, K., Wayne, A., & Yoon, K. S. (2007). State and local implementation of the No Child Left Behind Act, volume II—Teacher quality under NCLB: Interim report. Washington, D.C.: U.S. Department of Education, Office of Planning, Evaluation and Policy.

Borman, G.D., Hewes, G.M., Overmann, L.T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research, 73*, 125–230.

Boyd, W.L. (1978). The changing politics of curriculum policy making for American schools. *Review of Educational Research, 48*(4), 577-628.

Brophy, J.E. (1982). How teacher influence what is taught and learned in classrooms. *The Elementary School Journal, 83*(1), 1-13.

Brown, J., & Duguid, P. (1991). Organizational learning and communities of practice: Toward a unified view of working, learning, and innovation. *Organizational Science, 2*(1), 40-57.

Buckley, J. & Schneider, M. (2007). *Charter Schools: Hope or Hype?* Princeton, NJ: Princeton University Press.

Bryk, A.S. (2009). Support a science of performance improvement. *Phi Delta Kappan, 90*(8), 597-600.

Bryk, A.S., Sebring, P.B., Allensworth, E., Luppescu, S., & Easton, J.Q. (2010). *Organizing Schools for Improvement: Lessons from Chicago.* Chicago, IL: University of Chicago Press.

Carter, S.C. (2001). No Excuses: Lessons from 21 high-performing, high-poverty schools. Washington, DC: Heritage Foundation.

Center for Research on Educational Outcomes [CREDO]. (2009). *Multiple Choice: Charter School Performance in 16 States.* Retrieved June 29, 2009 from http://credo.stanford.edu.

Center on Reinventing Public Education and Mathematica. (2011). Charter-School Management Organizations: Diverse Strategies and Diverse Student Impacts.

Chetty, R., Friedman, J.N., and Rockoff, J.E. (2011). The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood. NBER Working Paper No. 17699.

Clark, M., Gleason, P., Tuttle, C.C., & Silverberg, M.K. (2011). Do charter schools improve student achievement? Evidence from a national randomized study. Mathematica Policy Research Working Paper.

Clotfelter, C., Ladd, H.F., & Vigdor, J.L. (2007). How and why do teacher credentials matter for student achievement? Working Paper: National Center for Analysis of Longitudinal Data in Education Research.

Cohen, D.K. (2011). *Teaching: Practice and its Predicaments*. Cambridge, MA: Harvard University Press.

Cohen, D.K. (2005). Professions of human improvement: Predicaments of teaching. In M. Nisan & O. Schremer (Eds.), *Educational Deliberations* (pp. 278-294). Jerusalem: Keter Publishers.

Cohen, D.K., Raudenbush, S., & Ball, D. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis, 25*(2), 1-24.

Cohen, D.K. and Spillane, J. (1992). Policy and practice: The relations between governance and instruction. In G. Grant (Ed.), *Review of Research in Education, 18*, 3-49.

Cohen, D.K., Gates, K., Glazer, J., Goldin, S., & Peurach, D. (forthcoming). *Improvement by Design.* Manuscript for book in progress.

Cook, S.D.N. & Brown, J. (1999). Bridging epistemologies: the generative dance between organizational knowledge and organizational knowing. *Organization Science, 10*(4), 381-400.

Cook, S.D.N. & Yanow, D. (1993). Culture and organizational learning. *Journal of Management Inquiry, 2*(4), 373–90.

Correnti, R. & Rowan, B. (2007). Opening up the black box: Literacy instruction in Schools Participating in Three Comprehensive School Reform Programs. *American Educational Research Journal, 44*(2), 298-338.

Curtis, R. (2011). Achievement First: Developing a teacher performance management system that recognizes excellence. Washington, DC: The Aspen Institute. Retrieved March, 2011 from http://www.aspeninstitute.org

De Mello, V.B. (2011). Mapping State Proficiency Standards Onto the NAEP Scales: Vcariation and Change in State Standards for Reading and Mathematics, 2005-2009. NCES: Washington, D.C.

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis, 24*(2), 81 - 112.

Doyle, D., & Han, J.G. (2012). Measuring teacher effectiveness: A look "under the hood" of teacher evaluation in 10 sites. New York: 50CAN; New Haven, CT: ConnCAN; and Chapel Hill, NC: Public Impact. Retrieved June, 2012 from http://www.conncan.org/learn/research/teachers/measuring-teacher-effectiveness

Duncan, A. (2009a, June 8). Remarks in Press Conference Call on Charter Schools and the Race to the Top Fund. Transcript retrieved June 30, 2009, from http://www.ed.gov/news/pressreleases/2009/06/06082009a-trans.doc

Duncan, A. (2009b, June 22). "Turning Around the Bottom 5 Percent." Remarks at the National Alliance for Public Charter Schools Conference. Transcript retrieved June 30, 2009 from http://www.publiccharters.org/DuncanKeynoteNCSC09

Education Sector. (2009). *Growing Pains: Scaling Up the Nation's Best Charter Schools.* Washington, DC: Education Sector. Retrieved November 24, 2009 from www.educationsector.org.

Elmore, R. (2004). *School Reform from the Inside Out: Policy, Practice, and Performance.* Cambridge, MA: Harvard Education Press.

Elmore, R.F. & Burney, D. (1997). Investing in Teacher Learning: Staff Development and Instructional Improvement in Community District #2, New York City. National Commission on Teaching and America's Future, the Consortium for Policy Research in Education.

Freeman, D. & Porter, A. (1989). Do Textbooks Dictate the Content of Mathematics Instruction in Elementary Schools? *American Educational Research Journal, 26*(3), 403-421.

Fullan, M. (2010). *All Systems Go: The Change Imperative for Whole System Reform.* Thousand Oaks, CA: Corwin.

Garet et al. (2008). The Impact of Two Professional Development Interventions on Early Reading Instruction and Achievement. Jessup, MD: Education Publications Center, IES.

Garet et al., (2011). Middle School Mathematics Professional Development Impact Study: Findings after the Second Year of Implementation. Jessup, MD: Education Publications Center, IES.

Gleason, P., Clark, M., Tuttle, C.C., and Dwoyer, E. (2010). The Evaluation of Charter School Impacts: Final Report." Princeton, NJ: Mathematica Policy Research.

Goldhaber, D.D., & Brewer, D.J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis, 22*(2), 129–146.

Gordon, R., Kane, T.J., and Staiger, D.O., (2006). Identifying Effective Teachers Using Performance on the Job. Hamilton Project Discussion Paper. Published by the Brookings Institution.

Hanushek, E.A. (2009). Teacher Deselection. In D. Goldhaber & J. Hannaway (Eds.), *Creating a New Teaching Profession.* (pp. 165-180). Washington, D.C.: Urban Institute Press.

Hanushek, E.A. & Rivkin, S.G. (2006). Teacher Quality. In E.A. Hanushek & F. Welch (Eds.), *Handbook of the Economics of Education, Volume 2* (pp. 1052-1078). Amsterdam: Elsevier.

Hill, P. T., Lake, R. J., Celio, M. B. (2002). *Charter schools and accountability in public education.* Washington, DC: Brookings Institution Press.

Hoxby, C. & Murarka, S. (2008). Methods of assessing achievement of students in charter schools. In M. Berends, M.G. Springer, and H.J. Walberg (Eds), *Charter School Outcomes*. New York, NY: Lawrence Earlbaum Associates.

Hoxby, C.M., Murarka, S., & Kang, J. (2009, August). *How New York City's charter schools affect achievement* (second report in series). Cambridge, MA: New York City Charter Schools Evaluation Project. Retrieved September 24, 2009 from http://www.nber.org.

Ingersoll, R. (2001). Teacher Turnover and Teacher Shortages: an Organizational Analysis. *American Educational Research Journal,* 38(3), 499-534.

Jerald, C. (2012). Movin' it and Improvin' it. Using both education strategies to increase teaching effectiveness. Center for American Progress.

Kane, T., Rockoff, J.E. and Staiger, D.O. (2006). What Does Certification Tell Us About Teacher Effectiveness? Evidence from New York City. NBER Working Paper #12155.

Ladson-Billings, G. (2006). Yes, but how do we do it? Practicing culturally relevant pedagogy. In J. Landsman and C. W Lewis (eds.) *White teachers/Diverse classrooms: A guide to building inclusive schools, promoting high expectations, and eliminating racism,* (pp. 29-42). Sterling, VA: Stylus Publishing.

Lake, R., Dusseault, B., Bowen, M., Demeritt, A., & Hill, P. (2010). *The National Study of Charter Management Organization (CMO) Effectiveness: Report on Interim Findings.* Retrieved July 1st, 2010 from http://www.crpe.org/cs/crpe/view/csr_pubs/344.

Lampert, M. (2001). *Teaching problems and the problems of teaching.* New Haven, CT: Yale University Press.

Lampert, M., & Graziani, F. (2009). Instructional activities as a tool for teachers' and teacher educators' learning. *Elementary School Journal,* 109(5), 491–509.

Lampert, M., Beasley, H., Ghousseini, H., Kazemi, E., & Franke, M. L. (2010). Using designed instructional activities to enable novices to manage ambitious mathematics teaching. In M. K. Stein & L. Kucan (Eds.), *Instructional explanations in the disciplines* (pp. 129-141). New York, NY: Springer.

Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. London: Sage.

Lipshitz, R., Popper, M., Friedman, V.J.. (2002). A Multifacet Model of Organizational Learning. *The Journal of Applied Behavioral Science, 38*(1), 78-98.

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

Maxwell, J. A. (2005). Qualitative research design: An interactive approach. (2nd Ed.). Thousand Oaks, CA: SAGE Publications.

Merseth, K.K. with Cooper, K., Roberts, J., Tieken, M.C., Valent, J., and Wynne, C. (2009). *Inside Urban Charter Schools.* Cambridge, MA: Harvard University Press.

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

Miles, M.B., & Huberman, A.M. (1984). Qualitative data analysis: A source book of new methods. Beverly Hills, CA: Sage.

Miles, K.H., Odden, A., Fermanich, M., & Archibald, S. (2004). Inside the black box of school district spending on professional development: Lessons from five urban districts. *Journal of Education Finance, 30*(1), 1-26.

National Alliance for Public Charter Schools (NAPCS). (2012). Dashboard. Accessed May 2012. http://dashboard.publiccharters.org/dashboard/home

NAPCS. (2011, October). A Growing Movement: America's largest charter school communities. Retrieved October, 2011 from www.publiccharters.org.

NewSchools Venture Fund. (2006, December 5). Charter management organizations: Toward scale with quality. San Francisco, CA: NewSchools Venture Fund.

NewSchools Venture Fund. (2010). AF Talent Review Overview, September 2009. Included in: A highly effective teacher in every classroom: Creating talent development systems that drive instructional excellence. San Francisco, CA: NewSchools Venture Fund. Retrieved April, 2010 from www.newschools.org.

Obama, B.H. (2009, March 10). [Remarks on a Complete and Competitive Education]. Speech presented to the Hispanic Chamber of Commerce, Washington. DC. Transcript retrieved June 30, 2009 from http://www.whitehouse.gov/the_press_office/Remarks-of-the-President-to-the-United-States-Hispanic-Chamber-of-Commerce/

O'Reilly, C., & Chatman, J. (1996). Culture as social control: Corporations, cults and commitment. In B. Staw & L. Cummings (Eds*.). Research in Organizational Behavior*: Vol. 18, Greenwich, CT: JAI Press, 157-200.

Orlikowski, W.J. (2002). Knowing in practice: Enacting a collective capability in distributed organizing. *Organization Science, 10*, 249–273.

Phillips, K. (2010). What Does "Highly Qualified" Mean for Student Achievement? Evaluating the Relationships between Teacher Quality Indicators and At-Risk Students' Mathematics and Reading Achievement Gains in First Grade. Elementary School Journal, v110 n4 p. 464-493.

Purkey, S. C., & Smith, M. S. (1983). Effective schools: A review. *Elementary School Journal, 83*(4), 427–452.

Raudenbush, S. (2009). The Brown legacy and the O'Connor challenge: Transforming schools in the image of children's potential. *Educational Researcher 38*(3):169-180.

Rivkin, S.G., Hanushek, E.A., & Kain, J.F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417-458.

Rosenberg, S. (2012). Scaffold, Safeguard, Shield, and Barrier: A Charter Network's Developing Infrastructure of Practice. Unpublished manuscript.

Rowan, B. (1990). Commitment and control: Alternative strategies for the organizational design of schools. *Review of Research in Education, 16*, 353–389.

Rowan, B., Camburn, E., and Barnes, C. (2004). Benefiting from comprehensive school reform: A review of research on CSR implementation. In C. Cross (Ed.), *Putting the pieces together: Lessons from comprehensive school reform research* (pp. 1-52). Washington, DC: National Clearinghouse for Comprehensive School Reform.

Rowan, B., & Correnti, R. (2009). Studying reading instruction with teacher logs: Lessons from a study of instructional improvement. *Educational Researcher,* 38(2), 120-131.

Rowan, B., Correnti, R., and Miller, R.J. (2002). What large-scale survey research tells us about teacher effects on student achievement: Insights from the Prospects Study of elementary schools. *Teachers College Record, 104*(8), 1525-1567.

Sanders, W.L., & Rivers, J. C. (1996). Cumulative and residual effects of teachers on future student academic achievement. Knoxville: University of Tennessee Value-Added Research and Assessment Center.

Schmidt, W.H. and R.T. Houang. (2007). Lack of focus in the mathematics curriculum: Symptom or cause? In Lessons learned: What international assessments tell us about math achievement, edited by Tom Loveless, 65-84. Washington, DC: Brookings Institution Press.

Schmidt, W.H., Houang, R., and Shakrani, S. (August, 2009). *International Lessons about National Standards*. Washington, DC: Thomas B. Fordham Institute.

Smith, M.S. & O'Day, J. (1990). Systemic School Reform. In *Politics of Education Yearbook, 1990* (pp. 233-267). New York: Taylor and Francis.

Sørensen. J.B. Sørensen. (2002). The Strength of Corporate Culture and the Reliability of Firm Performance. *Administrative Science Quarterly* 47: 70-91.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.

Stigler, J.W. & Thompson, B. (2009). Thoughts on creating, accumulating, and utilizing shareable knowledge to improve teaching. *Elementary School Journal*, 109, (5) 442-457.

Teh, B., McCullough, M. Gill, B. P. (2010). Student Achievement in New York City Middle Schools Affiliated with Achievement First and Uncommon Schools Mathematica Policy Research, Inc.

Tough, P. (2006, November 26). What it takes to make a student. New York Times.

Tuttle, C.C., B. Teh, I. Nichols-Barrer, B.P. Gill, and P. Gleason. (2010, June). Student Characteristics and Achievement in 22 KIPP Middle Schools. Washington, DC: Mathematica Policy Research.

Vogus, T.J., Sutcliffe, K.M., Weick, K.E. (2010) Doing No Harm: Enabling, Enacting, and Elaborating a Culture of Safety in Health Care. *Academy of Management Perspectives, 24*: 60-77.

Waller, W. (1932). *The sociology of teaching.* New York: John Wiley.

Wayne, A.J. & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research,* 73(1), 89-122.

Weick, K. (1976). Educational Organizations as Loosely Coupled Systems. *Administrative Science Quarterly, 21*(1), 1-19.

Weick, K., & Roberts, K.E. (1993). Collective mind in organizations: Heedful interrelating on flight decks. *Administrative Science Quarterly*, *38,* 57-381.

Wilson, S. (2008). *Success at scale in charter schooling.* American Enterprise Institute Working Paper. Retrieved May 18, 2009 from http://www.aei.org/outlook/100019.

Yin, R.K. (2009). *Case study research: Design and methods (4th ed.)* Thousand Oaks, CA: Sage.

CHAPTER FOUR

SCAFFOLD, SAFEGUARD, SHIELD AND BARRIER:
A CHARTER NETWORK'S DEVELOPING INFRASTRUCTURE OF PRACTICE

Establishing greater reliability in the quality of students' opportunities to learn in the United States must involve widespread improvements in teaching. Yet policymakers and practitioners who seek to improve teaching throughout a school system—whether a district, charter school network, state, or the nation—face at least three interdependent challenges related to the United States' historically decentralized educational system and its legacy, which is in turn related to the political system in which it is embedded (Cohen & Spillane, 1992; Meyer & Rowan, 1978; Weick & McDaniel, 1989).

First, within this largely incoherent system, would-be reformers typically control few of the instruments through which they might influence the aspects of schooling most likely to affect student achievement, or the "core of instructional practice" (Elmore, 2004).[66] These instruments include things like curricula, diagnostics and assessments, pedagogical routines, the organization of teachers' work, or the opportunities for learning afforded to them. Second, educators in the U.S. lack a unified professional knowledge base for practice or practice improvement that might inform the development and use of these instruments, organize them into coherent systems to guide educators' work, or provide a framework for building and sharing additional professional intelligence (Bryk, 2009; Hiebert, Gallimore, & Stigler, 2002). Finally, within this fragmented educational and political system, a strong tradition of local control in the governance and funding

---

[66] This "core" includes teachers' understanding of knowledge and learning, and how they are enacted in the classroom; the social and structural organization of teachers, students, and their work together; and ways of assessing and communicating about student learning.

of schools and the absence of shared instruments or professional knowledge has supported teachers' embrace of autonomy as a way to manage these uncertainties, which has resulted in a powerful idea teachers' work as "privatized, idiosyncratic practice" (Raudenbush, 2009, p. 172) in which skilled teaching is largely viewed as an individual trait. This pervasive set of beliefs may be responsible for reformers' reluctance to create or manipulate instruments to influence the instructional core, or for teachers' rejection of their attempts to do so.

In recent years, a prominent group of scholars in the field have proposed that the deliberate development of instructional systems or regimes (Bryk, 2009; Raudenbush, 2009; Wilson, 2008), or a more comprehensive "infrastructure" for educational practice (Cohen, 2011) may serve as a way to gain some purchase on each of these problems. According to Cohen (2011), an "infrastructure of practice," organizes "the extensive technical and professional affordances" that both "enable and sustain work" (p. 56). In other words, the resources that are part of an infrastructure of practice "are the frame around which skilled work is built" (p. 49). In education, an infrastructure of practice might combine curricular, educational, and social resources into a coherent system, which, if well-designed and used well, would support teachers in their work and its improvement; provide leaders with the tools needed to develop and manage teaching quality; and instantiate and enable educators' specialized knowledge and skill, serving as a dynamic repository professional learning. As teachers worked together with these communal tools in the pursuit of common aims, their efforts could foster a paradigm of teaching as "shared, systematic practice" (Raudenbush, 2009, p. 172), in which teachers share knowledge, skill, and responsibility to ensure their schools more reliably support students in achieving.

In a previous essay (Rosenberg, 2012), I profiled the teaching quality development and management strategies employed by Achievement First (AF), a high-performing charter

management organization (CMO) dedicated to providing its students, who tend to be poor and are almost all African American or Latino, "with the academic and character skills they need to graduate from top colleges, to succeed in a competitive world and to serve as the next generation of leaders for our communities" (AF, n.d.), and to doing so at the scale of an urban district to inspire broader educational reform. Based upon a case study of the network's efforts to define and organize to support high-quality teaching and learning across classrooms and schools in their rapidly growing network in 2009-2010, I argued that AF's leaders were leveraging their schools' relative autonomy as charters, the strengths of the individuals they hired, and their network approach to build an infrastructure of practice. Because of the absence of rich illustrations of what an emergent infrastructure of practice might look like in the U.S., I described the basic features of this infrastructure and their design, the ways they were expected to be used together, and their chief affordances as perceived by network and school leaders. These included both individualistic and systemic strategies for enhancing teacher and teaching quality such as:

Teacher recruitment and selection. AF developed a selective recruitment and selection process to identify and hire teachers who had been successful students themselves, who shared a commitment to mission of the organization, and who appeared to be willing and able to share responsibility for helping students to achieve these goals using AF's instructional tools.

Instructional framework and tools. In addition to a common commitment to the network's shared mission and core values, AF developed a set of tools and practices to guide teachers' work, students' learning opportunities, and the network's continuous improvement efforts. These included a clearly delineated cycle of effective instruction and set of ten essential principles and practices for instruction; a shared scope and sequence, set of interim assessments, and an online assessment platform, Athena, to support data analysis, reflection and planning.

Instructional leadership and organizational design. AF was organized to protect the work of instruction and instructional leadership by holding sacred and extending learning time for students and teachers and by carefully designing and distributing leadership responsibilities across positions and parts of the organization. Instructional leaders were hired in part because they were effective teachers themselves, often within AF. Like the teachers they worked with these leaders had access to a range of professional development within and across schools including individualized coaching and formal and informal opportunities to collaborate with colleagues. Leaders' work was also supported by a several formal talent management tools and practices.[67] AF's network structure, shared tools, and strong professional relationships provided the organization with opportunities to learn from "experiments" in each site.

Organizational culture. These three features of AF's infrastructure combined with, structured, and were structured by an organizational culture distinguished by a nearly fanatical performance orientation, a strong focus on individual and organizational growth, and a high degree of shared responsibility, mutual accountability, and trust. AF's organizational culture was deeply intertwined with the more tangible tools and practices that compose AF's infrastructure and supported their productive use in the specific contexts of educators' work.

I concluded the essay by highlighting two key challenges facing AF's leaders as they pursued their goals: first, the ways in which success in teaching and learning has been defined within the network and its surround, and second, the sustainability and potential for replication of their model.

---

[67] These included a professional growth plan (PGP) for professional evaluation, reflection, and goal setting; performance improvement plans (PIPs) for teachers who needed extra support, and that also supported transparency around termination; clear processes and programs to cultivate and recognize excellence in teaching and leadership; and school and network-wide talent reviews to assist in strategic planning for the network's talent needs.

Here, I draw upon the same case study data and methods to investigate in greater depth (a) how the elements of AF's infrastructure appeared to be enacted within the network and (b) what some of the critical opportunities and challenges associated with building and using an infrastructure of practice for teaching in the U.S. might be.[68] I found the infrastructure that AF was building functioned within the organization in ways that were so multi-faceted and remarkable that it requires several different perspectives to understand and appreciate them. First, as anticipated by Cohen (2011), Bryk (2009), and Raudenbush (2009), the AF infrastructure served as a *framework* and *scaffold* for individual and collective professional learning and practice, and as an *outcome* of these efforts. Second, AF's infrastructure also functioned as a *safety net*, or as a set of tools, systems, and practices that helped AF staff and students to reliably meet the goals the network set for them in part by fostering processes of "mindfulness" that characterize effective high reliability organizations (Weick & Sutcliffe, 2001). Third, I found that these features of AF's infrastructure created a *haven* for many of the young professionals it employed, sheltering their beliefs, work, and opportunities to learn from some of the disorder of the broader educational system. In each of these ways, AF's still nascent and uneven infrastructure provides a promising model for confronting several of the key challenges involved in improving teaching across an educational system.

Yet I also found that the extraordinary strengths of AF's organizational design, and the tools, practices, and culture that were part of it paradoxically created corresponding challenges for the organization and some of the individuals within it. This requires a fourth perspective, in which I describe the ways in which some AF staff members—especially more experienced teacher and school leaders—came to experience aspects of AF's infrastructure as an *impediment*

---

[68] For a description of these data and methods, and background on Achievement First, please see the "Data and Method" and "Achievement First Background" sections of Rosenberg (2012) where they are described in depth.

to individual and organizational growth. Several subjects also raised questions about how aspects of AF's infrastructure and culture might limit students' learning opportunities despite the many ways that they supported their success.

Below, I use each of these four themes in turn as lenses through which to examine (a) the labors and experiences of the teachers and leaders who worked on and within this emergent but rapidly developing infrastructure, and (b) the ways that they described the possibilities and challenges they encountered as they worked towards ensuring better and more reliable teaching quality and student learning throughout the network.

Lens I: Infrastructure as a Framework, Scaffold, and Outcome of Learning and Practice

One way to understand the infrastructure of practice that AF was building is as an evolving framework that organized, and scaffold that lent support and structure to, individual and collective professional learning and practice—and as a dynamic outcome of these efforts. A focus on individual teachers' learning was particularly important in AF because of the numbers of new teachers, or early career teachers new to the network, that AF employed as a consequence of their recruitment and selection model and extraordinary growth. In the 2009-2010 school year, 77 percent of teachers hired had been teaching for five or fewer years, and 36 percent were first year teachers (Curtis, 2011). But the same statistics are also indicative of the need for organizational learning and the development of systemic supports in AF: the network could not count on each new teacher to learn to teach and teach well on their own at the speed that was required for them to meet the organization's goals for student achievement, so they had to attempt to create conditions that would enable and sustain these novices' practice.

In fact, because AF hired so many new teachers, because there were few models for providing a reliably high-quality education for low-income or minority students within and across multiple schools, and because many of the elements of an infrastructure of practice were unavailable, rare, or unaligned in the broader U.S. context, AF and its employees were likely to have much to learn as they pursued their mission. This included not just knowledge for teaching, but also knowledge for corresponding and equally challenging work such as instructional leadership; building and using the tools, practices, structures and norms that would support this work; conducting essential political and operations work such as securing charters, funding, and facilities; and enacting all of these things successfully across contexts and time. AF had to provide its teachers and leaders with the resources they needed to practice while also finding ways to harness their learning for sustaining and improving the work of the organization as a whole. What is more, they had to manage to build this ever-expanding ship while sailing it.

Below, I provide a very brief introduction to some foundational ideas about learning, knowledge, and knowing in practice for individuals and organizations. I then offer several examples of their instantiation in AF concentrating largely on teachers and teaching in AF's elementary schools, which were the focus of the school-level investigations of this case study.

Individual and Organizational Knowledge and Learning in Theory

Defining, and attempting to describe the relationships among, individual and organizational knowledge, practice, and learning, is challenging due to the complexity of the concepts and controversy in the ways they have been conceptualized and studied across disciplines. For example, referring to "individual" learning may be somewhat misleading, because, like the concept of an infrastructure of practice, prominent theories of learning—particularly the learning

197

of complex practices like teaching—often underscore the interdependencies of individuals'

capabilities and the contexts in which they live, learn, and work. Instead of viewing learning as

the transmission of knowledge that may be possessed and transferred from one individual to

another, theorists including Lave & Wenger (1990) present learning as an ongoing process of

social construction.[69] Brown & Duguid (1991) offer a concise summary of this view in which

learners are cast as "construct[ing] their understanding out of a wide range of materials that

include ambient social and physical circumstances and the histories and social relations of the

people involved." In other words, "like a magpie with a nest, learning is built out of the materials

to hand and in relation to the structuring resources of local conditions. […] What is learned is

profoundly connected to the conditions in which it is learned" (p. 47). Indeed, they explain, from

this view learning is not about an individual's acquisition of expert knowledge or of "learning

about practice," but rather a process of enculturation, or of "*becoming* a practitioner" who is able

participate appropriately as part of a community (p. 48).

Whether or not organizations—and not just the individuals within them—are capable of

"learning" is subject to widespread debate since applying models of this human capability to an

organization appears to require reliance on either metaphor or an anthropomorphic view of

organizations, both of which obscure the processes through which a collective might learn.

Theorists have attempted to overcome this obstacle in different ways. For example, Cook &

Yanow (1993) argue that a "cognitive perspective on organizational learning" (p. 374) that either

equates organizational and individual learning or focuses on learning by key individuals and the

changes to an organization that they later initiate is useful but inadequate for understanding

learning *by* an organization. This latter concept requires establishing how an organization learns

---

[69] The social dimension of individual learning is one of the reasons that Berends, Boersma, & Weggeman (2003) draw upon Anthony Giddens' structuration theory to reject the "dualism of organization and individual in the study of organizations" (p. 1052).

to do something that may only be done by a group and in which the knowledge of how to do so

is embedded within the organization as a whole rather than a single person (e.g., the Celtics

playing basketball, an orchestra playing a symphony, or a workshop creating fine flutes). They

instead offer a "cultural perspective," in which organizations are cast as cultural rather than

cognitive entities. Arguing that "a culture is constituted, at least in part, from the intersubjective

meanings that its members express in their common practice through [artifacts, or] objects,

language, and acts" (p 386), they propose that learning by organizations occurs "through

activities involving cultural artifacts, and that learning, in turn, is understood to entail

organizations' acquiring, changing, or preserving their abilities to do what they know how to do"

(p. 386).

At the heart of these descriptions of individual and organizational learning is a focus on

learning to participate or to do something in the context of one's work, rather than just learning

about something. This notion raises questions about the nature of the knowledge that enables

individuals or groups to acquire, perform, or share competence. In their essay about different

forms of knowledge—tacit and explicit knowledge as they are possessed by individuals and

groups—and how they intersect in practice, Cook & Brown (1999) argue that what is known by

an individual or a group cannot be understood simply as *knowledge* that can be possessed, but

also must include *knowing* that is "part of practice" (p. 382).[70] They argue that the "generative

dance" between knowledge and knowing that occurs in practice can be an important source of

organizational innovation, in which "the creation of new knowledge and new ways of using

---

[70] By "practice," Cook and Brown (1999) specifically refer to "*the coordinated activities of individuals and groups in doing their 'real work' as it is informed by a particular organizational or group context,*" which they distinguish from both "behavior" and "action." They write: In the simplest case, if Vance's knee jerks, that is behavior. When Vance raps his knee with a physician's hammer to check his reflexes, it is behavior that has meaning, and thus is what we call action. If his physician raps his knee as part of an exam, it is practice. This is because the meaning of [the physician's] action comes from the organized contexts of her training and ongoing work in medicine (where it can draw on, contribute to, and be evaluated in the work of others in her field)" (pp. 386-387).

knowledge are possible." In other words, "knowing entails the use of knowledge as a tool in the interaction with the world" (p. 393). From this perspective, understanding the ways that competency might be disseminated would involve careful study of "how this essentially non-transferable or 'situated' dimension of knowledge and knowing, as elements of an organization's core competency, can be 'generated in' (rather than 'transferred to') other groups or organizations" (p. 398).

Orlikowski (2002) embraces Cook & Brown's (1999) concept of knowing though she argues that tacit knowledge is not distinct from knowing and action as they propose but is instead constituted through and inextricable from action (p. 251). She suggests that in addition to focusing on organizational knowledge, a complementary focus on "knowing as enacted in practice" has important implications for the ways in which we might understand competence to be accomplished, shared, and sustained within and across groups.

> A view of organizational knowing as an enacted capability suggests that core competencies or capabilities of the organization are not fixed or given properties, embodied in human resources, financial assets, technological artifacts, or infrastructural capital. Rather, they are constituted every day in the ongoing and situated practices of the organization's members. […] The conventional view is that competencies are stable properties of particular individuals or units that can be invoked as needed in different situations. Thus, when skillful performance does not ensue, commentators seek explanations in the failure of those properties ("human error") or breakdowns in the system ("equipment malfunction"). If, however, skillful performance is seen as an active accomplishment, its presence is not presumed and its absence is not sought in the failure of the parts. In contrast, when skillful performance is seen to lie in the dynamic engagement of individuals with the world at hand at a particular time and place, both its presence and absence are understood as emerging from situated practices. The focus then is on understanding the conditions (e.g., human, social, structural, financial, technological, infrastructural) under which skillful performance is more and less likely to be enacted. (p. 270)

Therefore, instead of simply seeking mechanisms for the identification and transfer of "best practices" throughout an organization or system, Orlikowski suggests we might imagine how to

develop "people's capacity to enact—in their own particular local situations—'useful' rather than 'best' practices" (p. 271). Because know-how and practice are "mutually constitutive," their distribution is neither a problem of transfer nor "a process of disembedding 'sticky' knowledge from one community of practice and embedding it in another" (p. 271). Instead, sharing knowing can be envisaged "as a process of enabling others to learn the practice that entails the 'knowing how.' It is a process of helping others develop the ability to enact—in a variety of contexts and conditions—the knowing in practice." (p. 271).[71]

Within AF, I found many formal and informal mechanisms to encourage both individual and organizational learning—whether the latter is understood as learning by individuals within organizations who make subsequent changes to the organization; as learning by the organization through its members' development of shared meanings and artifacts enacted in practice; as developing peoples' capacity to enact "useful" practices; or as creating the conditions likely to support skillful performance. While some processes within AF are focused more explicitly on either individual or organizational learning, in many cases these opportunities are organized in ways that have the potential to generate both.

Individual and Organizational Knowledge and Learning in AF

AF's infrastructure and culture helped to create conditions that supported individual and collective work and learning at the same time that this work and learning constructed or

---

[71] In a later essay, Orlikowski (2006) cautions that in addition to the social dimensions of knowledge and knowing, it is important to give thoughtful consideration to the role of the "material forms, artifacts, spaces, and infrastructures through which humans act" to understand knowledgeable practice (p. 460). Drawing upon "scaffolding," as a metaphor, she develops a notion of "knowing in practice as scaffolded – both culturally (e.g., through codes, language, norms) and materially (e.g., through physical objects, biological structures, special contexts, and technological artifacts)" (p. 462). This notion of knowing in practice as scaffolded through both cultural and material resources offers another useful way of understanding the notion of an infrastructure of practice and what its affordances might be for an individual's practice, and the reverse.

enhanced many features of this infrastructure and culture. First, AF's recruitment and selection process yielded large numbers of novice or early career teachers with a commitment to the network's mission and their own professional growth, strong critical thinking skills, and other competencies aligned with AF's core values and vision of effective instruction. These hires were widely understood to be central to the network's success; AF's leaders found that educators selected for these characteristics were generally motivated to improve, receptive to feedback, and willing to work collaboratively. They also tended to have serious time to invest in their work since they were frequently young and unencumbered by commitments to their own families. These qualities were important resources for an organization in which people were asked to be transparent about their practice and the outcomes of their teaching; to be open to feedback about their strengths and areas for growth; to reflect deeply on their own and others' practice; and to work closely with colleagues, students, and families as they engaged in and learned to do this work. These practices reinforced both individual dispositions and their cultural instantiation within or across schools.

Second, the instructional tools central to AF's infrastructure—common aims, the Cycle and Essentials, scope and sequence, interim assessments, Athena, and complementary protocols for data analysis and instructional planning—were also key resources for professional work and improvement because they helped to define the content of the work that was to be learned in and by AF. They served as frames around which shared language, meanings and practices could be built, and provided common assessments of student learning and standards for practice. These resources also structured and directed attention to common problems of practice, like how to best teach third grade students to write a response to literature essay; how to make sense of a set of interim assessment results and adjust instruction and other resources accordingly; or how to

engage in difficult conversations about performance with a struggling novice teacher and support her in becoming more proficient in a particular area of practice. Much of the shared work done on these problems occurred in the specific contexts in which they arose and had to be managed, and if not, people often brought artifacts from practice to inform the deliberations.

Finally, another significant element of AF's infrastructure, the design of the network and the ways in which AF organized the work of teaching, school leadership, and network leadership, also fostered individual and organizational learning. A fairly horizontal network structure combined with common instructional tools supported comparing, sharing, and learning within and across schools. The way instructional leadership in AF was designed promoted strong professional relationships among staff members within and across schools. These relationships were frequently described as being characterized by mutual respect, shared responsibility, and transparency about the successes and challenges people faced in their work. When these professional relationships spanned organizational boundaries, they also served as important interpersonal channels to connect individual or group knowledge and knowing via practice, creating conditions that were likely to support competent practice as AF defined it.

To illustrate the way that AF's infrastructure may work to support both individual and organizational learning, I provide several examples of AF's practices around data collection, analysis and planning; of the ways the network is organized, and organizes the work within it; and of how AF's infrastructure serves as an outcome and frame for some of these efforts.

Data collection, analysis, and planning

AF collected a wide variety of process and outcome data to be used for assessing and supporting the effectiveness of individual and collective practice. For example, in addition to student

achievement data, AF collected survey data from employees, students, and families about their performance at least once a year. After every event or workshop, attendees were asked to provide feedback on what was or was not useful about them. School performance data were captured by each school's Achievement First Report Card, which included a summary of achievement data and other metrics like teacher and student retention, student attendance, and parental satisfaction.

While in some cases these data were used for evaluative purposes (e.g., professional evaluations), data in AF were more typically employed in formatively to support ongoing learning and improvement. Organizational tools and practices helped to ensure these data were not only collected, but were also shared, analyzed, and integrated into subsequent changes in plans or policies throughout the network using carefully designed data platforms, data displays, and common practices and protocols supporting individual and organizational learning and a culture that encourages this learning.

These practices permeated the organization. For example, AFNS teams tried to exemplify their valued orientations toward practices of data collection, analysis, reflection, and transparency through formal inquiry into their performance using annual surveys. Sarah Coon, Director of Evaluation and Organization Development, explained that AFNS administered an "organizational healthy survey" once a year that focused on people's satisfaction in the network, and a "network support survey" to gauge people's perceptions of the support AFNS provided twice a year. In the fall, a bare bones version of the survey was given to principals, which included common questions the collegiality, helpfulness, responsiveness, and communication of every AFNS team. In the spring, everyone in the network took the survey, which included the same questions as the fall as well as questions about specific initiatives in the network, like the PGP process or AF's developing teacher career pathway program.

AFNS teams were expected to take these results very seriously. AF's Chief Operating Officer (COO), Maia Heyck-Merlin, explained that team leaders were responsible for going through the results with their teams and asking, "What are the quick fixes, things that we can just do; what are things we need to engage more deeply with principals on; and what are things we think are good ideas but we can't tackle yet and here's why." The results also informed AFNS teams' annual "first-class planning process" in which AF set priorities for the year, as did each team, which also assumed responsibility for some of them. Next, Coon explained, the teams "break those down into individual goals and tactics," drawing upon the network support survey results to do so. "If my team didn't do well this year with responsiveness, then in my first class plan I may have a specific goal around responsiveness related to that survey and really, some very clear tactics around how we're going to achieve that goal," she continued.

AF co-CEO and Superintendent Doug McCurry offered another example of how data might be used both to improve specific aspects of the network's work and to institutionalize the processes through which such improvements might be achieved. During network-wide principal cohort meetings held at least four times a year, he explained,

> One of the things we [repeatedly] do at our principal meetings is a focus on data. Each principal meeting we pick a different piece of data, and typically what we do is: here's the data, what is it telling us network-wide, where are we strong, where are we weak, where do we need to get better? And then honestly we look not for outlier lowers but for outlier highs.

To codify and share what they learned, principals of schools with "outlier highs" were asked to write up and discuss with the group what they believed had made them so in the principal cohort meetings.[72] However, to McCurry, equally important was the opportunity to reinforce values and processes around the practice of using data to reflect and improve:

---

[72] Principals and other coaches typically reported finding these meetings valuable in part because of these opportunities to hear more about others' approaches. As one principal pointed out, their busy schedules hindered

> The idea is we're trying to institutionalize this practice of, let's regularly look at data, let's confront the brutal facts, and let's learn from the people that are doing a better job. The next time [we meet] we do that with the AF report card, then we have the org health data at the middle of the year. And so we look for the trends, what are we seeing? Alright, network-wide we're a little low on this particular metric, although this one school is high, what are they doing, hear from them, and [try] to institutionalize that practice so that that's just the norm.

At the same time, McCurry emphasized "everything has context" that needed to be carefully applied to the analysis of the data, which was another valued norm within the network.

School leaders reported employing similar approaches to modeling these orientations and practices for their staff. For example, one principal explained that the feedback he received from his staff about his performance from the school leader survey was useful for his professional growth but also gave him an opportunity to provide an example of an ideal response to feedback on performance for his staff.

> When we get these survey results, which we take really seriously—I mean, I treat it for me the same way as we treat data day for kids in achievement: incredibly transparent and really—not only do I share all the data, but the way that I share it, and the way that our deans share it and our grade level chairs share it is – these are the skills that you've assessed [me as] proficient at and I'm going to keep on doing them; these are the skills that you think that I'm in need of growth and I feel very accountable to that, just as if we were looking at your class' reading scores and deciding with kids are in the green, which kids are in the yellow, and which standards are mastered and which aren't. (Interview44)

By modeling reflection, humility, and transparency as they attempted to make sense of and act upon the data they collected, school and network leaders encouraged the use of data in similar ways throughout the organization.

Resources like carefully scheduled time to work with colleagues during common prep periods during the work day, or on special occasions like data days, encouraged the use of these data in these more formative ways as well, as did the individual dispositions toward reflection

---

collaboration, but at these meetings, "if you hear something of interest that you want to follow up, then you can either go to that school or talk to that principal to follow up," (Interview19).

that the organization prioritized in the hiring process and critical instructional tools. So did

directed practice on this practice: the Essentials directed teachers to collect daily "exit slips"

aligned with their lesson aims to assess student's burgeoning understandings (or lack thereof)

and were expected to adjust planning immediately according to the results. Interim assessment

results every six weeks were immediately fed into AF's internally developed online assessment

platform, Athena, which allowed professionals working throughout the organization to have

access to all students' IA results and to generate a variety of reports that that could be used in

conjunction with other data and observations to plan for instruction, to identify students,

teachers, or school leaders in need of support, or to plan network-wide interventions or initiatives

(see the section of this essay about grade level teams and Rosenberg, 2012 for more about data-

driven instructional planning and protocols within the network).

The tools and practices described in this section provide examples of organizational

learning whether or not it is understood as resulting from an individual learning and making

subsequent changes that influence collective practice; as the development of shared meanings

and practices; or as creating conditions in which people are likely to practice skillfully. However,

the likelihood of this learning being productive increases because of key features of the

network's organization, which I discuss below.

Network organization

Shared instructional tools, the data they generated, and maintaining a staff with a common set of

commitments and mindsets allowed AF to leverage some of the strengths of being organized as a

network for learning. First, developing and using of a common infrastructure across the

network's schools permitted educators and network leaders to make sense of their own

performance outcomes, to motivate improvement, and to reach out to others with similar

struggles or who had been more successful in their work. McCurry explained,

> When we just had one school we gave interim assessments and a teacher would do [poorly and say], "Oh the scope and sequence is really hard, etc." Well, if you have four or five middle schools in the same geography and everyone is taking interim assessment number two in the sixth grade and one school does really well, and a couple schools in the middle, and one really poorly, the school that does really poorly can't just say, "Oh well you know it was too hard, it was this, it was that," right? They should look and say, "Oh wow, we're not doing [well], what are those other schools doing?" And then it allows us to be able to see who are the teachers across the network that are really making really strong growth, and so could we learn from them and share? Are there schools that are doing a particularly great job on a particular standard? And without everyone [taking the same test at the same time] it's very, very challenging for us to do that.

Second, by sharing instructional tools, AFNS was able to look for data trends across the network

to help its leaders decide when and where to intervene, and how. If, for example, they noticed

that second grade math was a low point across the network, AF's Chief Information Officer

Harris Ferrell suggested they were able to ask:

> Are there things that we should be doing as a network to help improve instruction in second grade math, and do we have the models of what excellence looks like, do we have the right scope and sequence, are there gaps in learning, are there materials that could be provided, is there better training that we should be doing? Because if these are the challenges that we're seeing across our schools then we shouldn't ask each school to try to solve this problem on its own, let's engage the schools together and let's see if we can solve it together.

Common use of these critical frameworks across the network also allowed AFNS to coordinate

cross-school professional development so that it was timely and more useful.

Finally, use of a common set of instructional tools by a committed staff, and strong

relationships between AFNS and the schools—fostered in part by their relationship of mutual

accountability structured formally by the charter management agreements that bound them and

partly by an organizational culture that emphasized shared responsibility and work—was part of

what individuals within AF believed allowed for greater flexibility in other ways at the school

level. Explained an academic dean:

> Here's where [AFNS] won't budge: you will have five IA cycles, every school will take the same IA, … [elementary schools] will give a midyear and an end-of-year reading assessment, and the [Fountas and Pinnell Benchmark Assessment System] is what we're using. So there are certain things that they don't budge on because they want us all to be able to compare with each other and say, this is how we're doing together, we are a unit. What you do in between those things is up to your school. But […] each school has a superintendent that watches over them and they come to the school and they observe these things. And if you want to make a change you explain it and you tell them why you think that this is good and we're going to try it. […] So I do feel like they're very open to hearing [the school wants to try something new] as long as you understand that you are taking on this project, it's a pilot at your school, and if it goes well you might be sharing it with other people; if it goes bad, then you're willing to say, all right, this didn't happen the way that we wanted it to happen so we're not going to do it anymore […] the goal is to get kids to college, so as long as we keep on that path, how we get there is [flexible, but] needs to [include] solid research based education tools. (Interview21)

These comments illustrate several of the ways of AF's infrastructure and its use facilitated

learning within and by AF. First, the framework provided by the infrastructure of practice

allowed for professional flexibility within schools and classrooms because of the structure and

oversight provided by frequent monitoring of student performance through IA and other

assessment results, and of school leader or teacher performance through coaching and other close

work with colleagues.[73] This arrangement enabled local innovation or improvisation in the form

---

[73] An AFNS leader gave an example of an AFNS initiative that highlights some of the benefits of this flexibility and the utility of having both IA results and coaches to limit the possible liabilities of this flexibility: AFNS had created a small number of reading units for 5th and 6th grades to serve as resources for teachers and coaches. She observed that teachers tended to use these units in one of three ways: "There's the teacher who takes the unit and just follows it lockstep and doesn't do enough thinking on their own before implementing, and then they have challenges. There are teachers who really look at it critically and are [think], this is the backbone but I'm really going to use this to make it specific for my kids and be careful about the text I select and change some aims around, […] but I'm going use it as a great starter. And then there are people who just completely ignore it, the units, and are doing their own thing completely. You know, the best case scenario is two; We want you to be adapting them to your class and figuring out what will work for your kids, so we expect variability, because there needs to be but they are good, solid units and you shouldn't totally toss them out, but you also shouldn't follow them lockstep." She added that when people tend to simply toss them out, it generally happens because "the owner of reading at a school is not clearly defined. People are just making individual choices and no one is holding them accountable. I think that is one way that the IA's do help." She said she hoped that AFNS was helping deans and principals to hold people

of resources or practices that were adopted, developed, or adapted in particular settings to make their way into the organization through individuals or groups of individuals; offered safeguards to minimize their potential liabilities if they were not effective; and provided channels through which others within the organization, and the organization itself, could learn.[74] This flexibility combined with a commitment to shared tools and partnerships also provided a way for people within the network, who were hired and valued in part because of their industriousness, to exercise ownership over their work and to contribute to the organization, potentially enhancing their commitment to AF.

Organization of work and learning within the network: overlapping professional networks
Professional or "talent" development was a major priority within AF. Some of this development was organized as somewhat traditional professional development sessions and workshops.[75] Teachers' formal learning began before the school year with a 3-week intensive staff training, two weeks of which were run by the network with input from schools and were targeted to teachers new to AF, before spending a week at the school site. This new-staff training introduced teachers to critical instructional tools and to AF's culture.[76] In addition, the biannual AF-wide PD days involve a series of workshop sessions led by both internal and external personnel.

accountable for using the materials or for their results, "and if they are making up materials on their own that aren't effective, then pointing them in the direction of good materials" (Interview17).

[74] The adoption of F&P across elementary schools is an example of learning of this kind in which a small number of schools were successful with a particular program that was then extended across the network.

[75] Network leaders thought of these as simply one small part of their professional development efforts. A network leader noted that too frequently in education, she observed people "worshipping at the golden calf of PD [as it is typically done], […] like, 'PD will solve our problems. People just need more training.' And that's not really it, you need more practice doing the right things, practicing the right things, and hopefully PD gives them the tools to do the right things and there needs to be follow up" (Interview 17).

[76] COO Heyck-Merlin explained, "[New staff training] is both the culture building part of things, you know: Are we a strong network, is there high enthusiasm, energy, [do] the people feel connected? All those things. And then just the nuts and bolts of our Essentials of Instruction, building everybody's vision of excellence to get there."

Friday afternoon PD sessions at each school frequently included whole-school or small-group workshops led by school or network leaders.[77]

However, key crucibles for individual and collective learning in AF were what I have come to think of as overlapping professional networks. Overlapping professional networks (OPNs) are *groups that serve as sites for sustained formal and informal collaboration between two or more teachers, school leaders, or network leaders and that produce, refine, and draw upon key elements of an infrastructure of practice in their work together, creating conditions supportive of individual and collective learning.* Within schools, these overlapping professional networks included one-on-one coaching relationships, K-2 teacher pairs, grade level or content area teams, administrative teams, and the entire school staff. Across schools, professional networks include principals and their AFNS coaches; cohorts of principals, academic deans, grade level or content chairs and leadership fellows, each facilitated by AFNS leaders; and other standing or temporary AFNS teams. By definition, individuals are always members of more than one professional network that often spans organizational boundaries, hence the "overlapping" moniker.[78]

Working within these OPNs had several important affordances for teacher, school leader, and network leader learning and practice. These included (a) providing practitioners with sustained and structured access to colleagues and to what they know and can do, including a coach specifically charged with their professional development and support, and (b) making

---

[77] These small group sessions were sometimes differentiated according to teachers' needs. For example, school leaders or teachers with particular strengths did things like organize brief workshops on a particular topic or practice that teachers can choose between. At other times, the staff organized themselves into temporary working groups that examined and provided feedback about a teacher's practice via video protocols, engaged in joint lesson planning or curriculum mapping and writing, or worked on writing a shared professional literature review.

[78] The idea of an OPN is very similar to the concept of a "professional learning community organized around a specific instructional system" that Bryk (2009) suggests may be a crucial mechanism for "the social organization of improvement" for individual and collective learning of complex work like teaching (p. 599). Yet I gave the concept a new name to (a) emphasize their overlapping character, and (b) to allow myself the room to develop the concept without misrepresenting Bryk's notion.

work more manageable by providing practitioners with key resources for their work vis-à-vis the infrastructure and colleagues with whom to share or distribute their labors. Together, these things were reported to enable practitioners to (c) accomplish much more than they might have alone, and to (d) feel more capable and supported in their work. They also provided leaders with opportunities to (e) introduce or model desired practices, and (f) with a way of staying connected to the work of those they lead, which allowed them to better identify the strengths and weaknesses of particular practitioners, of network policies and tools as they were used in practice, and then with opportunities to better support individual or organizational improvement.

Without a shared infrastructure of practice—recruitment and selection processes that provided teachers with academically successful like-minded colleagues committed to similar aims; instructional tools that gave content to the work and a way to assess progress toward these goals; and an organizational design that fostered mutual accountability and shared responsibility while designating structured time and leadership roles that created opportunities to work together—the existence of and productive work of these OPNs would be much less likely. Together, AF's mutually constitutive infrastructure of practice and culture provided staff members working in OPNs with critical resources for their learning and practice, and with a framework for organizing individual and collective practice improvement efforts. The network's infrastructure and culture are also evolving *outcomes* of this work that they help to structure since many OPNs worked specifically on developing or adapting tools or practices that compose the infrastructure. Two tangible examples are the working groups that drafted the Essentials or that refined or rewrote the IAs. A less tangible example included the ways in which collective work over time helped to institutionalize shared norms or culture around things like data practices or engagement with students' families.

Below, I provide more detailed examples of the ways people described their work in three kinds of OPNs—coaching dyads, grade level teams, and cohort-based OPNs—to illustrate more concretely the ways in which OPNs appeared to offer important affordances for individual and collective work and learning in AF, and the ways in which AF's infrastructure of practice and culture helped to constitute and was constituted by their efforts.

Coaching. Coaching was the primary and most formal relationship directed at supporting individual learning. In AF, coaching created sustained opportunities for teachers, school leaders, or network leaders to work with a colleague charged with helping them to improve their practice and students' achievement. In 2009-2010, AF was in its second year of formally developing a school-based coaching model in which every teacher in the network met with a coach—usually a principal, principal-in-residence, or academic dean, though sometimes a more experienced teacher—once every week or two to work continuously toward mastery of a learning goal outlined in a dynamic teacher learning plan. Coaches' work with teachers was largely structured by the Essentials (see Appendix D). By using the Essentials as a framework for coaching, AF's leaders hoped to help teachers to improve their practice in ways that were generalizable across many different kinds of lessons. The content of coaches' work with teachers was also influenced by AF's scope and sequence, the IAs, and the data they generated. In addition, during the 2009-2010 school year, AFNS leaders were in the process of developing a resource called "Coaching the Essentials," a flexible "scope and sequence" for the coaching of teachers in the Essentials that moved from basic management to work on core instructional strategies to more "advanced" topics such as deeper engagement, rigor, or character development and that provided coaches with resources linked to each of these areas (AF, 2010b).

Coaches typically described their coaching in ways that suggested a sustained engagement in practitioner development with the goal of improved student achievement and in ways that roughly corresponded with the network's coaching to mastery strategy though they were afforded wide latitude within that frame. In addition, coaches highlighted some of the relational work they did to support teachers with any challenges that they might be experiencing in addition to the explicit and more sustained instructional focus of the interactions. For example, when asked to describe her work as a coach, an academic dean explained that she met with most of the teachers in her coaching portfolio every other week. She sent out a meeting agenda based on the teacher's learning plan and her classroom observations in advance, asking the teacher if he or she wanted to add anything to discuss, such as "a difficult conversation with a parent, or [if] she got test scores back and they were really disappointing." As for the meetings themselves:

> I'd say my best coaching meetings might start with: I observed you teaching guided reading yesterday, let me give you some feedback on what I saw, and run through all the great things that the teacher did, and then one or two "big rocks," areas to grow. And then we'll say, let me look at your guided reading plans for next week, and let's see how we can make sure that you're accounting for rigorous questioning, or something like that in the plans. And then we'll sit there and we will work on rigorous questioning, and maybe we'll bring up a video of a teacher who really had rigorous questioning and see what techniques they used, and put it in there. We'll do that, and then we make commitments, so they'll say they'll email their updated plans by Friday. And then I'll say okay, and I'm going to come and observe you at least once next week to see how it goes. And the cycle begins again. (Interview28)

If the coach found that some teachers needed even more intensive support, he or she might also engage in more intensive co-teaching or co-planning models, or in real-time coaching to support teachers during instruction.[79]

---

[79] For example, a principal who had launched her school the previous year with 10 out of 14 teachers being novices discussed some of the benefits of adding a co-teaching component for some of her teachers in her coaching portfolio: "It was a way for me to be in all classrooms and schedule out time so I would know what's going on in all the classrooms, and to be able to do a coaching model where it's not just, every two weeks we come and sit down and have this conversation where I can support you on your one or two particular things. Co-teaching with you and

The success of this model depends in part on the strength of the coaches and in part on how receptive people are to being coached in these ways. Teachers and leaders reported that AF teachers were typically extremely reflective and eager to learn from feedback. Indeed, although some veteran teachers interviewed raised concerns about the execution of the coaching model, the teachers I spoke with overwhelmingly valued their experience with coaching. One elementary school teacher explained,

> What coaching [is like when it works as it should] is one of the best things, I think, one of the things that draws me right into AF. You have someone to watch you teach and pinpoint something and say, all right, this is what we're going to work on. It's like you're a student yourself learning how to master your craft. We went into this job knowing we wanted to be amazing teachers for our students, so who better to help us than teachers who have been teaching for a little bit longer, [who] can help you overcome things that they probably went through themselves? (Interview46)

When I asked a grade level leader who was also a coach whether or not people were generally receptive to her feedback, she responded:

> I think one of the best messages that AF sends is that feedback is a gift. They say it all the time and people embrace it and it's true. It's very, very, very true. […] I mean, when I get feedback I take it and I think about it and I reflect on it and I apply it where I feel I can. And I'll give pushback if I don't feel it's good feedback and nobody gets angry. And as a coach, […] I don't know if this is just me being able to gently give it or if it's just well received, but I think I've never really had an experience where people were angry about the feedback. I don't know, people—there's just a culture of wanting to grow here. (Interview41)

---

co-planning, it just has a stronger impact and I think a longer lasting effect […]. Right now [I'm coteaching with a teacher] doing grammar and reading comprehension, so we just talk about, what do we see as the kids are struggling with and then come up with an action plan, here's how we explicitly need to teach it. And she takes the lead on it and I'll just interject [including while she is teaching…]. I'll just say, Ms. So-and-So, do we maybe want to do this, or how do we think about this, or is this a question? I'll just chime in via a question that I would be asking. The goal with that is that it doesn't seem that I'm correcting her in any way whatsoever, it's [that] we're tag-teaming a conversation and helping" (Interview18). In addition, in 2009-2010, some AFNS instructional leaders had begun to work with Lee Canter or experiment with his real-time coaching approach, particularly with respect to novice teachers' classroom management struggles.

Both comments indicate some of the ways in which the qualities that factored heavily into teacher selection, such as a growth orientation and commitment to excellence, strengthen and were strengthened by the structured opportunities they had to learn within the network.

Teachers were not the only people with coaches in the network. School leaders also had coaches (e.g., principals were coached by a superintendent or regional superintendent, principals coach school leaders and usually some teachers, etc.), which helped to cultivate a culture of continuous improvement. As with the coaching of teachers, providing leaders with coaches was expected to create sustained opportunities for practitioners to work with a strong thought partner on personalized professional development in the specific contexts of their work. And as with the coaching of teachers, one-on-one coaching of leaders gave their coaches a way to become deeply familiar with their performance in ways that provided different and more immediate information than simple outcome or process metrics and allowed them to stay strongly connected to the particular challenges involved in school leadership within AF. In addition, the cascading coaching model provided coaches with opportunities to provide examples of strong coaching practices for leaders who are coaches themselves.[80]

Grade level teams. The second OPN I profile here is the grade level team. For elementary school teachers, grade-level teams were often discussed as critical sites through which

---

[80] Regional Superintendent Marc Michaelson, explained that by coaching principals, he could support them in their work by helping them with very different and specific areas of need, but also by providing them with strong examples of coaching. He explained, "I have been developing learning development plans for my own principals. I think it's important to model what we want to see our leaders doing in their schools in the way we approach our growth with them. […] Just to give two examples, one principal I'm working on leadership coaching actually, so he's doing a lot of distributive leadership this year and so I'm helping him to work on his coaching of his coaches and his team. And then with another, this principal just has more challenges around personal effectiveness, so some follow through issues and personal organization and things like that, so putting together a plan of how the two of us can tackle that together. So it's really scoping it out over a course of four weeks or so to figure out exactly what the interventions are, what are we going to do together that's going to build that skill. For the first [principal], it may be that we plan out some of the coaching meetings together and then I observe him doing them and give him some feedback afterwards. […] The second one, I've got a lot of ideas but I'm trying to get this principal invested in this work on personal effectiveness, so it's going to be a lot of conversations and then helping him to set up both his Outlook and his email systems just to be more effective, and time management is going to be a big piece of it."

experience, expertise, and workload were distributed among members of the group, though in

some elementary schools and in AF's middle and high schools content teams were often equally

or more important.[81] On data days between IA cycles every six weeks, teachers within a grade

level who taught the same subjects worked together with tools from AF (analysis and planning

protocols, Athena, the scope and sequence, the next set of IA assessments, etc.) to analyze IA

and other instructional data; to identify the standards students had mastered or needed to review;

to create or modify an aims sequence aligned to the scope and sequence and any other curricula

they were using; and to create a general plan for what needed to be accomplished in the next IA

cycle and how students and instructional resources might be organized to meet these goals.

Usually a grade level chair facilitated, though often an academic dean, principal, or principal-in-

residence attended these meetings or regularly consulted with them.[82]

AF's Chief Information Officer Harris Ferrell suggested that AF's understanding of the

benefits of these collective labors had developed considerably over time. While discussing the

way AFNS encouraged principals to structure data days at their schools, he explained,

> We've evolved [from a time when data-driven planning was more independent]
> and realized that the more team planning that happens around data, the more
> powerful. So rather than having each second grade teacher make their own data-
> driven plan […] the second grade team will look at the data together across the
> grade level. There will be some specifics in terms of your own students'
> performance, but in terms of grappling with what is the misunderstanding around
> these standards, there's a lot more commonality, and by having that dialogue
> across teachers you really get to a deeper level of understanding and you also

---

[81] In grades K-2 in AF, another structure supporting access to more experienced colleagues was the co-teaching model that assigned two teachers to a classroom with about 30 children. AF's leaders believed this design to be good for students' learning because it allowed for more time spent in small groups focused on specific groups' needs, but also had the potential to support teachers' learning. As one principal explained, "We usually will put our first year teachers, most of which are TFA, in with a co-teacher, so that there are two in the classroom. […] So we pair up new people with our strongest people. And then they might be there for a year and then we'll pair them up with someone else" (Interview18). This afforded new teachers near-constant access to a more experienced colleague who in turn had the opportunity to reflect on practice with the novice.

[82] Grade or content leaders also typically meet with other grade or content leaders and school leaders weekly, allowing them to share ideas and information with and from their teams, they are important liaisons between their team and the school for issues of this kind.

mitigate for some of the newer teachers who might not be as familiar with the content, or [have a way for] the veteran teachers to impart more of their wisdom as they do their analysis. And so a lot of the analysis happening at the grade team level, or content team level, has become a powerful way to help teachers improve their own classroom level planning.

Ferrell's comments point to the potential of OPNs to create opportunities for discourse that allowed for and drew upon the context-bound nature of their practice at the same time that divergent perspectives, expertise, and other resources could be brought to bear on individual and collective understandings and practices.

In addition to working together on data days, the teams typically met once or twice a week during scheduled shared planning times during the school day. These meetings often included discussion of logistics like fieldtrip plans or "cultural" issues relevant across the grade, but the main focus was usually on adapting and elaborating the plans created during data days to develop more detailed lesson plans. Some teams planned together, while others divided these responsibilities among individuals who "owned" planning for a specific area and then met to share, discuss, and refine plans as a group.

Many teachers reported that shared instructional planning during data day or weekly meetings allowed them to distribute the intensive workload and intellectual labors of teaching in an AF school, resulting in a more sustainable and better quality performance. For example, one teacher explained how much she valued the ways she and her grade level team divided responsibility for planning for different subject areas among them. She reported that she really liked, and liked working with, the other teachers and the academic dean who coached each of them. Without working together in these ways, she stated, "I don't know that I could still work here because it's so much work. Like I was saying with my first year before we were working together, I just don't know how anybody could do it, with [the work schedule] being 7:00 to 4:00

and needing all of that to be planned meticulously" (Interview32). Another teacher on the same team ageed: "It's really helpful because it's difficult to come up with quality plans for four different subjects, and so it really allows for the plans to be a lot better quality." She explained that the learning opportunities available through this process extended beyond those embodied in the plans themselves because "it isn't like somebody creates them and then they hand them out." Instead, lesson plans were usually distributed in advance of their weekly meetings and people were expected to arrive having reviewed them, "so if we have any questions we can ask the point person to clarify anything. Or if there are modifications that we need to [make for our specific group], we talk about that at that time too," (Interview34). Given these potential affordances, school leaders frequently reported being strategic in the ways they divided ongoing instructional planning, assigning more experienced teachers responsibility for planning in more challenging instructional areas and giving more straightforward planning, like that for "Textual analysis" (TA), the time in which students practiced basic reading comprehension, to less experienced teachers—sometimes with a gradual transition to novices doing more of that planning in consultation with their grade level chair or coaches.

Another mechanism for sharing expertise in grade level teams was through direct feedback on the quality of the plans, which grade-level leaders and coaches also frequently provided to the teachers with whom they worked. For example, a grade level leader who coached a small number of the teachers in her team explained that though she did not teach writing that year, she collected and provided written or verbal feedback on her team's writing plans:

> I might pull up the [plan] and actually write comments in red. […] Usually my suggestions have to do with either pacing or asking more rigorous questions. I might throw in a rigorous question just so that they can see what it looks like. I might say, your exit slip doesn't match your aim, go back and look at it. […] But because we're such a small community it's not that hard to just grab someone and be like, okay, this isn't going to fly because [of this] here. (Interview41)

Elementary school teachers and instructional leaders commonly reported similar formal and informal interactions about instructional plans or students.

Cohorts. As with grade level or content teams for teachers, other important opportunities for leaders within the network to learn came from sustained work with colleagues with the same or similar roles (e.g., groups of principals, or AFNS directors), or with overlapping responsibilities (e.g., the AFNS curriculum and professional development team). These cohort-based OPNs had affordances for both practitioners and for their leaders. They provide time for practitioners to work closely with colleagues, and through them, with ideas, resources, and others' experience and expertise to support them in developing competence in the core practices in their work; with opportunities to establish common meanings and norms across members of the cohort, as with establishing shared understandings about the meaning of the Essentials; as a way to work together on common problems of practice; as a way to roll out or get feedback on network initiatives; and as a way to foster relationships that translate into continued collaboration.

Some work in cohorts was quite deliberately planned and facilitated by AFNS leaders who worked to develop several mechanisms through which teachers might learn with and from their colleagues about core, common problems of practice. For example, Sara Keenan, AF's Director of Leadership Development, reported that in twice-yearly formal workshops for coaches, the AFNS leadership development team worked with the coaches on a case study that modeled the process of selecting a learning goal for a teacher, as well as the initial coaching meeting in which the coach and teacher developed the teacher-learning plan.[83] In one of these sessions, "We showed some video and data of a teacher, we talked together [about what] might

---

[83] In addition to workshops focused on coaching teachers to mastery, coaches were also able to choose to attend sessions later in the day focused on the learning goals most relevant to teachers in their portfolio, such exit tickets, student engagement, or increasing rigor.

be the learning goals, and then we picked one, and then Doug [McCurry] and I actually

fishbowled what that might look like and people talked about it," she explained. Similarly, in

November, AFNS had a coach share a case study focusing on her ongoing work with a teacher.

The coach had identified a learning goal and worked toward mastery of this goal with the teacher,

and shared an account of this work with her peers. Then, the coach cohort watched video of the

teachers' instruction and discussed the extent to which the teacher had mastered the goal, and

what the coach's next steps might be. Keenan reported that this was "really effective" because

"everyone got to see this one coach talk about working with this teacher on tone, which is a

really hard thing to work on," and because they were able to discuss and give feedback on the

coaching process in the context of an authentic example of a learning goal, of a teacher's practice,

and of the coach's practice.

In addition to network-wide cohort meetings, within geographic regions, principals of

schools serving similar grades hosted the others and at least one member of AFNS' leadership

team to engage in similar work in the context of their schools. Marc Michaelson, a former

Amistad Academy teacher, then AF middle school principal who had just become a regional

superintendent in 2010, offered a description of the way the Connecticut middle school principal

cohort tended to operate:

> We do inter-visitations between the different schools, and so at each of those, all
> the principals observe at one school and give that principal feedback on their
> school and on strengths of instruction and growth areas. I think that really impacts
> and influences the planning that those principals do in their prioritization of
> building their schools. They just trust each other a lot. There are a lot of smart
> people in that cohort and they just—I think it's a very powerful, powerful thing.

During these inter-visitations, the group also engaged in "dilemma discussions," in which the

host principal presented a particular dilemma or set of dilemmas to work on with the group. With

an AFNS staff member in attendance as well, if common dilemmas emerged that might warrant

network-wide intervention or support, there was a clear and immediate channel for that to be communicated with AFNS.

Work in these OPNs with shared instructional tools and aims led people to ruminate on their collaboration in ways that seemed to reflect two of AF's core values: "team and family," and "many minds, one mission." A principal maintained:

> There are so many resources for how to do things on a very foundational level, so many templates and calendars and a really good effort to reach out to all the schools and just say "this principal is doing this," that kind of sharing of materials, which is extraordinary. I mean, I had never been in an organization that shared in that way. […] It's like we're all here together, we're all working together, we all want to lift each other up and hold each other up and in the spirit of that I'm going to share all of this information. (Interview49)

An academic dean expressed similar sentiments, "I think that's the beauty of the Achievement First network compared to some single site charter schools. […] We have a network of a whole lot of people and over time, you get to know who is really good at what, and so you can pull and ask questions and learn from each other" (Interview28).

Outcome of and framework for learning

In addition to serving as a framework and scaffold for individual and collective learning AF's infrastructure and culture were also dynamic "outcomes" of the network's learning. For example, the Essentials—which provided a critical framework that structured for much of the organization's talent development efforts and instruction—were also products of the organization's learning. The document itself was drafted internally in 2008 to explicitly articulate and codify a network-wide vision of instructional excellence. In creating this framework, AF personnel drew upon research on teaching; similar efforts by peer organizations like TFA; and AF leaders' and teachers' observations of the practices that they found were most effective in raising student achievement across grade levels or content areas. However,

developing the ideas that came to be codified in the Essentials reflected a much more extended process of collective learning.

Co-CEO and Superintendent McCurry traced the evolution of this vision of teaching excellence in broad strokes during his interview, suggesting that its development was largely the result of an ongoing series of discoveries about what was necessary but not sufficient for supporting strong teaching and learning across classrooms and schools—even with the strong "talent" that survived AF's selection process. He explained that Amistad Academy, and then AF, first focused on making sure that their teachers established a strong classroom culture and engaged in basic lesson planning. However, they soon found these features of their model were necessary but insufficient to ensure teachers' effectiveness with students so they turned their attention to strong goal setting and the use of formative assessments to gauge their progress toward them. These additions too proved to be necessary but not sufficient, and AF added an emphasis about the core lesson cycle and instructional elements to their view of effective practice.[84] In his assessment, in the spring of 2010 the "vast majority" of AF classrooms were "pretty solid on both of those dimensions, meaning they have strong enough classroom management and an understanding of the core lesson cycle." Despite AF's strong achievement results across the network, AF discovered that more still was necessary to sustain instruction that would support students' achievement and future success, so the organization was working to learn how to better support teachers on all of the above in addition to more refined aspects of instruction within that structure, such as engaging 100% of students, establishing extremely clear expectations for written and oral student work, rigorous questioning, and ensuring that the

---

[84] McCurry explained that in AF this core involved having clear aims and exit tickets daily, along with a strong mini-lesson aligned with the aims and enough independent work time to allow students many opportunities to practice what they are learning.

students were doing the intellectual work of the classroom—all had come to be part of the Essentials.[85]

Other outcomes of AF's learning were the result of the refinement of existing instruments. AF leaders adjusted the Essentials to reflect new ideas or evidence about effective instructional strategies as they emerged. Indeed, the "Note" under the Modeling/Guided practice section of the Essentials is the result of some teachers, school leaders, and network leaders asking for more specific guidance about how to incorporate a more inquiry-based, or constructivist, approach to teaching within the Essentials framework.[86] Another example involves the IAs. Each year, AFNS leaders analyzed the extent to which IA results were predictive of success on the state assessments and made adjustments to the scope and sequence and assessments in to improve their alignment. In addition, McCurry explained,

> We have a process where folks from the school come together to meet with network support to look at our scope and sequences and interim assessments every year to say what modifications […] we need to make. And then […] team CPD drafts interim assessments based on it, then schools get feedback on them before they go live. We need to do even tighter job on that but that is something we're really pushing to, [and] my hope is those get more and more rigorous as time goes on.

An academic dean provided an example of this kind of back and forth with AFNS. She and her teachers found that their students were performing extremely well on the last IA of second grade,

---

[85] He explained that while these were things AF had always believed in, and that featured in the instruction of strong teachers within the network, "it's just a matter of emphasis and getting [these additional practices] across the board."

[86] Despite AF's strong, dominant organizational culture and Essentials-based vision of instructional excellence, this point about this added note about exceptions to the I/we/you lesson structure points to some internal debate within AF about some aspects of the Essentials. Several teachers, school leaders, and network leaders made reference to this debate. For example, one elementary school academic dean explained that for the most part he found there to be a "shared vision" of strong teaching among teachers. However, he noted an interesting tension between some of the principles instantiated in the Essentials and some teachers' ideas of good teaching, particularly with a group of teachers who had attended Bank Street. Their orientation tended to be a bit more constructivist than the typical AF approach. He reported he thought it was a positive thing for the network and that his conversations with these teachers were pushing his own learning as someone whose entire teaching career, and first year as an academic dean, had taken place within AF. He worried that going too far in the direction of constructivism could mean that people sometimes lost focus on the outcome of the lesson—what kids are able to do—and instead focused too much on the fact that students were talking and engaged, but he thought a balance was good, and that he was engaged in a particularly useful "productive conflict" about this with one of the teachers he coached (Interview29).

but then the same students would do terribly on the first IAs of third grade. She explained that the first year they noticed this, they were frustrated because they did not believe the drop off reflected what was happening in their classrooms. When it happened again the next year, she and her team decided to contact AFNS. According to McCurry, AFNS did see this drop across many schools in the network, and in consultation with them decided "our second grade is probably not rigorous enough in terms of our scope and sequence and interim assessments," so they worked to make the second grade scope and sequence more rigorous to better prepare students for third grade, and so "teachers aren't getting false positives on how their kids are doing." The academic dean reported being pleased they reached out to AFNS. "They totally revamped the end of second grade, beginning of third grade IAs. So it worked, right? We just had to open our mouths" (Interview28). As this academic dean's comments suggest, the network's responsiveness to feedback sometimes increased practitioners' allegiance to AF.[87]

As people in the network adopted tools (such as the F&P continuum, or the STEP assessment system) or created them (as with the Essentials and the IAs), they not only refined the tools themselves as a result of what they learned from using them, but they also sometimes created additional supports for their use, such as protocols detailing processes and practices that others found effective, or offering annotated examples of the ways tools had been used well or poorly throughout the network. Some of these, like detailed, month-by-month protocols for operations teams at each school carefully detailing what must be accomplished by the team and when with aligned resources were understood to require adaptation to the particular circumstances of the school, but were required; others, such as the data analysis and planning

---

[87] Discussing the way that she had observed the network grow and change over time in response to the experiences, ideas, and expertise that she and others brought to or developed within the network, an AF principal explained, "that's one of the things [that left her] feeling like I was part of this network in such a deep way, because I felt like, 'I'm seeing my feedback in action in pockets'" of AF's work, particularly toward the beginning of the network's expansion (Interview49).

protocols provided to teachers for use between IA cycles were strongly encouraged by AFNS and sometimes required by principals; still others were simply available on the AF server for those who wanted them, like sample PGPs and conversations about them of varying quality.

Despite these accomplishments, many leaders and teachers observed that AF still had enormous amounts of progress to make when it came to one of the areas that they believed mattered most to AF's ability to be consistently effective: network-wide sharing about more detailed aspects of instruction. Regional superintendent Michaelson explained that while there were "piecemeal" attempts to share instructional resources across schools, "I would say [instructional sharing] is one of the next realms for Achievement First to tackle." He continued,

> Because now we do have really strong teachers in pretty much every subject level at every grade somewhere in the network, and so if we can get them to build out their unit plans and their yearly plans, their unit plans and their lesson plans, with materials for a year, that would give people some resources to build from because I would like to see some more consistency in instruction across the network.[88]

An academic dean agreed, explaining that if, the following year, a teacher went off and spent 12 hours writing a persuasive writing unit again when they did the same thing this year, he was going to "scream" (Interview29). Instead, he said he wanted his school and the network to get to the point where they can draw upon and improve on what happened before—at the same time that he thought these resources had to be accompanied by thoughtful and sustained learning opportunities for teachers. In fact, AF was beginning to take on some of this kind of work in the spring of 2010 (for more on this topic, see the "Possibilities" section of this essay). Despite these areas for growth, across the network, people reported feeling a lot of excitement about, and pride in, what they were accomplishing through building and sharing resources even if they still had a

---

[88] He also reported that at that time, the network was grappling with how to design a AFNS-level leadership position that was something like a "Director of Sharing" focused on sharing within and outside of the network to help to accelerate some of this work.

great deal of progress left to make. A teacher who also served as a grade level chair and coach, expressed this common view:

> I think network-wide, [and] at this school, we have a fantastic plan in place and we are still young, as both an organization and a school, and really excited as we develop it. And as individuals at different schools think of different ways to do things better and begin to share all those network-wide—I mean the sharing/collaboration piece, whether it's teacher-to-teacher, in a school, across the schools, coach-to-coach, same thing, dean-to-dean—it's just a phenomenal thing. I think it's really at the heart of our success. (Interview45)

This kind of excitement about what the organization was building and accomplishing seemed to be a useful foil for the network-wide focus on continuous improvement, which could in a different context contribute to a feeling of consistent failure rather than consistent improvement.

…

These examples illustrate some of the most important ways that AF was working to support individual and organizational learning during the 2009-2010 school year. By engaging in similar data collection, analysis, and planning processes at all levels of the organization, individual practitioners were able to reflect upon and adjust their own practice, or to make changes to organizational resources or practices that had the potential to influence the organization more broadly. The ways that the network was organized and the culture AF cultivated helped to support these practices throughout the organization. As it was used within the OPNs it helped to structure, AF's infrastructure provided opportunities for distributed expertise across and within groups. Indeed, to the extent that learning by an organization is cast as it is by Cook & Yanow (1993) as (a) the acquisition of knowledge that is embedded in and can only be done by a group and (b) cultural, occurring through collective work with cultural "artifacts," (or embedded objects, language, and acts) to acquire, change, or preserve organizations' ability "to do what they know how to do" (p. 386), the descriptions of co-teaching, grade-level or content teams,

227

leadership cohorts, and coaching dyads provided above may be understood not just as supports for individual learning and practice, but also of organizational learning. In other words, these and other OPNs can be understood as sites in which its members' interactions are shaped by artifacts that are both resources for and outcomes of the organization's shared work (i.e., the Essentials) around which they continually worked to establish and reconstitute shared meanings and practices, helping AF to acquire, evolve, or preserve the organization's ability to support their students in achieving across subject areas, classrooms, and time. From another perspective, as they constituted and were constituted by AF's infrastructure and culture, OPNs appeared to offer resources to create the conditions that enabled teachers, school leaders, and network leaders to learn to practice in ways that AF defined as skilled (Orlikowski, 2002), or to scaffold knowledgeability in practice (Orlikowski, 2006).

Lens II: Infrastructure of Practice as a Safety Net for Quality Outcomes

A second way to conceptualize of AF's infrastructure of practice and corresponding organizational culture is as a set of evolving tools, resources, and practices that provide a safety net (a) to prevent individuals or groups of individuals within the network—students, teachers, or leaders—from struggling without being noticed, and therefore, (b) to avoid, contain, or work around failure to support student learning. In other words, though it was still developing, AF's infrastructure and culture supported AF's efforts to realize what has historically been an elusive goal in the United States: to ensure reliably strong performance among low-income and minority students and their teachers across the rapidly expanding network's classrooms and schools.

In the organizational studies literature, the term "high reliability organization" is typically reserved for organizations like nuclear power plants or naval aircraft carriers that manage to operate successfully in the face of trying conditions and great uncertainty (Sørensen, 2002;

228

Weick, Sutcliffe, & Obstfeld, 2008/1999; Weick & Sutcliffe, 2001). Although AF may initially

appear to have little in common with these organizations, like classic HROs, the organization

needed to find ways to manage frequent uncertainty and unexpected events—in classroom

instruction, in supporting teacher learning, in building instructional technologies, in handling the

AFNS-school relationships and the political and fiscal environments in which they operate,

etc.—and to manage these dynamic interactions repeatedly, over time, as the network grew, with

people and contexts that were always changing. And although the repercussions of failures in

performance were perhaps not on the magnitude of those life and death stakes often faced by

traditional HROs, there were high-stakes for AF if they did not help their students to perform

well: ethically, in terms of children's social and economic trajectories, and politically, in terms of

securing support from families, funders, and the entities that authorize AF's charters. Indeed,

insights about the processes that foster classic HROs' effectiveness provide a useful framework

for illustrating some of the ways in which AF's infrastructure might be understood as a (still-

developing) safety net for students and teachers.

Here, I draw upon the scholarship of Weick and Sutcliffe (2001) in which they identify

five processes of "mindfulness" that support HROs' reliable performance for this analysis.[89]

Observing that "it is impossible to manage any organization solely by means of mindless control

systems that depend on rules, plans, routines, stable categories, and fixed criteria for correct

performance" because "no one knows enough to design such a system so that it can cope with a

---

[89] Weick, Sutcliffe, and Obstfeld (2008/1999) explain their focus on these cognitive processes in HROs by offering a definition of reliability that challenges the conventional wisdom about it. They write, "For a system to remain reliable, it must somehow handle unforeseen situations in ways that forestall unintended consequences. This is where previous definitions of reliability are misleading. They equate reliability with a lack of variance in performance. The problem is, unvarying procedures can't handle what they didn't anticipate" (p. 35). Instead, they argue HROs tend to involve variation in activity or performance, but "stability in the cognitive processes that make sense of this activity" (p. 35). They go on to propose that the five processes that contribute to the mindfulness of the organization therefore foster the organization's ability to discover and manage unexpected events to achieve reliable outcomes (p. 37).

dynamic environment," they argue that "designers who want to hold dynamic systems together have to organize in ways that evoke mindful work" (p. 49). They propose five core processes that cultivate this "mindfulness" and support the effective performance exhibited by HROs under the challenging conditions they typically face: a preoccupation with failure; a reluctance to accept simplifications; sensitivity to operations; a commitment to resilience; and deference to expertise. Below, I describe each of these processes before providing a few examples of the ways in which they are articulated through AF's infrastructure and culture.

Preoccupation with failure

Rather than focusing on success and becoming complacent, HROs are *preoccupied with failure*, and learn from failures (or near misses) that do occur. Weick and Sutcliffe (2001) write,

> Effective HROs both encourage the reporting of errors and make the most of any failures that are reported. In fact, they tend to view any failure, no matter how small, as a window on the system as a whole. They view any lapse as a signal of possible weakness in other portions of the system. This is a very different approach from most organizations, which tend to localize failures and view them as specific, independent problems. (p. 56)

These dispositions are "evident in frequent incident reviews, the reporting of error no matter how inconsequential, and employees' obsession with the liabilities of success" (p. 54).

AF did celebrate success, but staff members were also extremely concerned with the possibility of failing to support students' learning. In addition to simply wanting to guard against the political and fiscal problems that might be associated with lapses in organizational performance as measured by student outcomes, a large part of this concern was derived from their collective commitments to prepare their traditionally disadvantaged students to succeed in college and beyond, and to prove that it can be done at the scale of an urban district. For example, in discussing her intense commitment to her students and workload, an AF principal

described her experience pitching the school to parents of rising Kindergartners and first graders

the year before they opened their doors.

> Basically we're selling a school to parents, and […] they believed the story I was telling them about this school that didn't exist. And when I think about the kind of trust that parents and families put into us, on that level, I mean, I'm like, I'm not working hard enough. You know, that's the kind of stuff that keeps me up at night because you promise these people, you look them in the eye and you tell them, I will take your child to that place, I will give your child choice, I will give them access to education that you may not have ever had or think that you could have. (Interview49)

Similarly powerful commitments and feelings of responsibility—qualities central to AF's

selection and hiring process—provided many AFNS leaders, school leaders with the motivation

to attend to individual instances of failure.

AF's organization as a K-12 network created a kind of longitudinal accountability system

through which they received feedback on their success in preparing students at each grade as

they progress to the next—and especially as early cohorts of AF high school graduates began to

report back from their experiences in college. This organization helped to encourage AF's

leaders to engage in systemic analyses and problem solving for situations that might in other

educational contexts be treated as isolated events. For example, one AFNS leader provided two

examples of how the current state assessment systems failed to establish content or proficiency

standards that would adequately reflect students' preparation for future academic work:

> In New York, a 3 is passing and for the math, […] it's a 1, 2, 3, 4 scale and if you get to 3 you're proficient, and 4 is advanced, but 3 doesn't really mean much. We find the kids who got a 3 […] on the 8th grade Pre-Algebra course, and then [take] 9th grade Algebra I are not doing very well. So 3 doesn't really mean much, so we really have to think more and more about pushing more kids to the advanced level [in earlier grades]. And I think on the reading side, kids' writing is not really assessed [by the state], their writing is pretty terrible. […] When our 9th graders have gotten to high school, we have had to do a lot of triage work on writing. It's just they have a long, long way to go. When I think about fundamental skill for success in college and life, writing is just huge. (Interview17)

This type of feedback allowed educators and network leaders to identify problems like this and to work backward to support teachers and students in earlier grades in an attempt to avoid getting to a place where this kind of "triage" was necessary.

Weick and Sutcliffe (2001) note that to encourage error reporting so failure or near-failures may be identified and analyzed, it is important that people feel safe in doing so; when people are rewarded for reporting errors or mistakes, it can "strengthen an organizationwide culture that values reporting" (p. 55). Indeed, AF's dedication to their mission and focus on continuous improvement fostered a culture that valued identifying and reflecting upon individual and collective strengths and weaknesses. For example, the senior network leader quoted above described AF as having a "supportive feedback driven culture," in which, "you get more feedback than you could possibly imagine." She elaborated,

> Someone pointed out to me recently, we were at a director day event for learning for directors, and someone said, I've never been in room where your boss is also in the room and you're admitting all of your failures and inabilities to manage, and everyone's just spewing it out. It's like, I'm having this problem, I'm not doing this right. [And she said] I've never been in a place where that happened, you know? It's usually [that] when you're in front of your manager, you're like, well this is what I did right and this is what I did right and, really selling yourself. So I think that the humility gene in all of us is pretty strong, which makes it a nice place to work, but I think sometimes people are more critical of themselves than perhaps they should be. But yeah, it's a really wonderful place to work. (Interview17)

School leaders and teachers tended to report having similar relationships with their colleagues. Reflecting on his work with the Connecticut middle school principal cohort while he was a principal, Regional Superintendent Michaelson explained the processes by which the group had come to value their work together and "to like and trust each other." Their first cohort meeting was at Elm City, his previous school, at a difficult moment in the school's existence:

> We were really struggling with our culture and knew we needed to make some changes and some things had just declined that we needed to work on. And so I

really opened up and said, hey, I want your honest feedback. I have a feeling it's going to smart, but this is the way we push our school and this is the way we're going to grow. And so people really gave it to me hard [laughs]. I got what I asked for. But I think that it actually set a really positive tone [and established that] all of us can put our schools out there and we can feel like, look, we're here to help each other, so whatever we get from this is—it's something that we're going to be able to use and feel good about, pushing—because everybody wants to make their school great. There's nobody who doesn't want to do that. And so I just think that, like that was a piece of it. And, you know, likewise, people who put some pretty tough issues on the floor for dilemmas got very honest input. Everyone valued it. Everyone learned a lot.

Over time, Michaelson explained, the group developed a great deal of mutual respect. He added, "I don't think there's any simple formula for this, but I think people have to get to a place where they can be comfortable really putting themselves and their schools out there, and we've gotten to that place."

Reluctance to simplify interpretations

Weick and Sutcliffe (2001) also found that effective HROs are "*reluctant to accept simplifications* [and…] take deliberate steps to create more complete and nuanced pictures" (p. 11). This helps them to avoid developing expectations that can blind people to disconfirming evidence that might help them to anticipate unexpected problems, and often involves creating opportunities for respectful interaction "among people who have diverse expectations" (p. 61), since "it takes variety to control variety" (p. 62).

The ways that OPNs operated in AF often fostered the kind of "constant interaction" among individuals "who have diverse expectations" (p. 60) that Weick and Sutcliffe maintain helps to cultivate organizational mindfulness. While each member of AF belonged to more than one OPN, some positions, like the regional superintendents, principals, and academic deans whose responsibilities included serving as liaisons from classrooms and grade levels to the schools as a whole, or from schools to the network as a whole, played particularly powerful roles

in this regard. The diverse perspectives these individuals brought to their work in OPNs meant that they were uniquely positioned to help the group to identify and manage challenges they faced in their collective work. And as the descriptions of the director day event and the middle school principals' cohort in the previous section suggest, people typically reported a sense of collective trust in the intentions of their colleagues that facilitated the kinds of interactions that allowed them to not only report performance problems but also to take them seriously and to engage in collective problem solving about them when they did.

Similarly, despite AF's very serious commitment to using data, and especially student achievement data, to assess their progress toward their short and long-term goals, network and school leaders frequently reported the importance of balancing these metrics with others. For example, in talking about the importance of contextualizing data, McCurry gave the example of looking at the data of a school that was previously struggling with that school's new principal: "[I don't say,] 'Oh my God, you're so far below these other schools, what are you doing?' [I say], 'I see the slope of the line and it's pretty steep and you're rapidly catching up, and you're still below but wow, that's really awesome, way to go.'" Nonetheless, a minority of interviewees shared concerns about the ways they saw the focus on achievement as it is measured by state exams as distorting other goals for student learning (see Rosenberg, 2012 for more on this topic).

Sensitivity to operations

The third characteristic of mindfulness involves demonstrating *sensitivity to operations*, or establishing ways for some members of the HROs "integrate information about operations and performance into a single picture of the overall situation and its operational status" (Weick & Sutcliffe, 2001, p. 65). By paying "constant attention to real-time information," through

"frequent operations meetings, widely disseminated operational measures of performance, and nearly continuous face-to-face interaction" (p. 65), HROs are able to notice small problems early and to provide them with "undivided, wide-spread attention" (p. 63) and to prevent them from developing into larger problems.

AF's infrastructure offered several mechanisms that contributed to the network's sensitivity to operations. For example, the ways in which AF organized the work of teaching and leadership across the network and the prevalence of OPNs created opportunities for frequent face-to-face interaction among diverse members of the network characterized by mutual respect and trust, which encouraged feedback being shared and used for individual and collective improvement. AF's multiple, formal data systems provided useful information for professionals working within these OPNs who sought an overall picture of network, school, grade level, and classroom performance on a variety of metrics, and to use such information for subsequent planning and inquiry cycles. These data systems included AF's student information system run by Infinite Campus that included information on things like student grades and daily attendance updated daily; the AF balanced report card that provided an overview of school performance on a variety of academic and non-academic outcomes like attendance averages and student and teacher retention; and Athena, an online assessment platform that organized student achievement data from interim assessments administered roughly every six weeks and generated a variety of data displays and reports designed to be useful for people at all levels of the system, but especially for teachers. In addition, network support, organizational health, and school leader surveys provided critical feedback on the extent to which professionals throughout the network felt supported in their work and believe AFNS to be addressing school and classroom needs appropriately.

Together with these more formal mechanisms for data collection about performance, the nested organization of work and leadership in OPNs promoted a layered and distributed sense of responsibility for teacher and student learning that provided additional information for leaders as they made sense of the network's ongoing areas of success or trouble. Discussing the role of the grade level chair, one principal explained,

> That's the key paradigm shift of growing to be a leader, right? When you are able to take responsibility for something that is not directly your own, and when you start to understand all of the indirect ways in which you can help that other teacher help her kids do better and therefore be in some way ultimately responsible for how another group of kids do.

After explaining that he was ultimately responsible for student achievement in his school, but that he delegated a great deal of direct responsibility for grade level performance to his academic deans and grade level chairs, he added:

> I'm uncomfortable with the total image of like a flowchart down, [since] I think there are so many other opportunities outside of the grade level structure for every teacher to offer a ton of leadership and a ton of ideas, and I think that everything we do tries to be very responsive to what the staff is asking for and wants, whether it's informally, whether it's our […] taking the temperature of how people are doing and what they need, or if it's being really responsive to things like our annual organizational health survey and our organizational school leadership survey. So in some ways I'm at the top of that, but in some ways, if I were making a flowchart, there'd be lots of little arrows coming out from every teacher. (Interview44).

The way this work was organized also helped to ensure that lapses in performance on the part of teachers or students were noticed and addressed relatively early. In this and a previous essay (Rosenberg, 2012), I described several of the ways that struggling teachers were identified by their coaches, and how formal PGP or PIP plans detailing their areas of struggle, plans for their improvement, and the support they needed may be marshaled when typical coaching cycles were inadequate; I also described the ways that teachers whose failings were egregious or who did not make adequate improvements may be terminated, sometimes before the end of the year.

In the case of students, AF had a coherent set of principles and strategies underlying their approach to identifying and working to support struggling students, but the ways that they were articulated in elementary schools in the network varied somewhat depending on student need and the ways in which principals decided to organize the school's budget and student support positions. Across the network, AF tended to adopt the philosophy while some children had biologically-based learning disabilities and needed instructional support specific to those disabilities, many children could avoid special education if they had access to early and intensive instructional interventions. Depending on the resources available at their schools and students' needs, beyond the usual differentiation and support in instruction during the school day, after school, or sometimes on weekends included things like targeted individualized instruction with a Reading Recovery teacher or small group instruction with their own or another teacher or school leader using a response to intervention program or a similar supplementary intervention system. These interventions also offered teachers and school leaders more regular feedback on student progress, and frequently another professional's perspective on students' strengths or challenges as a learner.

Commitment to resilience

The first three of the five processes of mindfulness in effective HROs identified by Weick and Sutcliffe are largely aimed at anticipating and avoiding detrimental unexpected events, or at preventing them from becoming larger or more dangerous problems. However, because "no system is perfect" and mistakes do occur, the authors argue that HROs also demonstrate a *commitment to resilience*, meaning "they work to develop knowledge, capability for swift

feedback, faster learning, speed and accuracy of communication, experiential variety, skill and recombination of existing response repertoires, and comfort with improvisation" (p. 70).

The design of AF's infrastructure of practice exhibited an organizational commitment to resilience in a number of ways. First, AF's recruitment and selection process sought to identify people who demonstrated resilience through other experiences and in the face of strong feedback about their practice and areas for improvement. These individual dispositions, as they were exercised repeatedly in AF, helped to constitute and reflect a culture of resilience that mirrored one of their core values, doing "whatever it takes" to support student achievement. Second, because of their commitments to their students' learning, AF's teachers and leaders were almost constantly working to assess and improve their resources and the ways they supported their use which required them to be "willing to begin treating an anomaly even before they have made a full diagnosis […] in the belief that their action will enable them to gain experience and a clearer picture of what they are treating" (p. 69). In COO Heyck-Merlin's view, AF might have been described as "constantly improving but not jumping from fad to fad. I think we're pretty good at incrementally improving." People at all levels of the organization frequently described their improvement efforts as involving a careful balance between urgency because of their need to better serve their current students, and a more incremental approach to improvement because of the realities of building a system and their desire to learn about what worked or did not work about a new initiative as they developed it.

Weick and Sutcliffe (2001) suggest that organizational resilience is also supported by the presence of uncommitted resources that provide an organization with the reserves necessary to respond to respond flexibly to difficult or unexpected events without jeopardizing performance. This is another way that AF differed from many traditional educational systems. Particularly

important areas of "slack" existed within school leadership, with their roles and responsibilities organized so that there were many more people—e.g., academic deans, a director of operations, a dean of students, grade level leaders, a principal-in-residence, and a principal—to share the work of running a school than is typical for U.S. schools. McCurry explained that this was intentional:

> It's my general belief that schools are typically [drastically] undermanaged organizations. The ratio of manager to employee is radically off in schools […]. Most literature will tell you [that] you can successfully manage somewhere between six and ten people. […] And often times [schools have] one principal and forty people […]. And so even at a full-grown school [in AF], it's not exactly there but if we have a coach who has a portfolio of somewhere between eight and twelve people it's bigger than what we'd ideally like but it's small enough that we can really intensely work with some folks.

When a teacher (and their students) needed intensive support, when a teacher was fired mid-year and someone needed to step in immediately to ensure kids did not miss out on instruction, or when an extra instructor was needed to enable a particular grade level to divide students into appropriate small groups for instruction in a given IA cycle, these "extra" personnel existed to make these things happen.

In addition, AFNS personnel took on some of the essential tasks of running a school or served as thought partners for school personnel, taking some of the burden from them as worked to create successful schools. A school leader who was going to be opening a new AF school the following year described AFNS as "a phenomenal resource." This person continued,

> I couldn't—I'm sure I could, but I would not want to—open a school without them. I mean, there are so many things that they do to support our work and just so much thinking, and I think that's the best part of it for me, is while I'm bogged down in the everyday of, this child needs this, and this parent has this concern, and this teacher has this challenge—and just always trying to keep my head above water—there are people at the network who are thinking, what will make coaching even more effective? Or what really are the things we should be looking for in a teacher, or how can we streamline our human resources policies? Just, everything; there are smart people thinking about it and then bringing it back. And they're incredibly open to feedback and it never feels like this is a directive from the network, you must do this. It's, we've been thinking, we've developed these

tools, and what do you think of them, and what is your feedback, what would you like to use? And sometimes it's, how can we help you use this, but always in a good way. I think they do a great job pushing schools to be better and supporting their work. (Interview39)

These comments suggest another feature of AF's culture that contributes to a final process of mindfulness: deference to expertise.

Deference to expertise

The last of the processes that Weick and Sutcliffe (2001) identify as contributing to mindfulness in HROs demonstrate *deference to expertise*, appealing to the "people with the most expertise, regardless of their rank" (p. 16), particularly as unexpected problems occur. Weick and Sutcliffe (2001) explain, "by blending a hierarchical decision structure with a specialist decision structure HROs recognize—and operationalize—a principle that often escapes decision makers at critical moments: Expertise and experience are usually more important than rank" (p. 74). Since in HROs "decisions have to be made quickly and accurately," those who are closest to the problems are often "empowered to make important decisions and are held highly accountable for those decisions" (p. 75). However, when an event is particularly unusual or high-stakes, they may seek the support of higher-ups in making decisions, or to "push decision making back up the hierarchy" (p. 75). AF's network structure; distributed leadership model within and across schools; close working relationships characterized by mutual respect and trust developed through shared work in OPNs; and a culture that encouraged people to ask for help from colleagues when they needed it all contributed to AF's practices around this process of mindfulness.

…

As enacted by AF, the network's infrastructure and culture cultivated these five processes of mindfulness, which provided them with important resources for identifying both network-wide

areas for improvement and individuals' struggles so that they could marshal the resources necessary to circumvent, or intervene to contain, possible organizational "failures"—ultimately providing a "safety net" for professionals and students and greater reliability in student learning and other organizational outcomes. Instead of relying on static, standardized scripts or their faithful "implementation" in schools and classrooms to achieve reliable outcomes, working toward stability in these processes of mindfulness helped to focus the network and the people working within it on learning from and for their work, and allowed them room to adapt and innovate in the particular contexts of their practice while still relentlessly pursuing consistently strong outcomes. In other words these processes of mindfulness helped to strengthen the ways in which AF's infrastructure acted as a powerful but nimble frame and scaffold for individual and organizational learning, distinguishing AF from many educational systems in the U.S. and contributing to the experience of many within the network of their infrastructure of practice as a professional haven.

Lens III: Infrastructure of Practice as a Professional Haven

A third way to view the infrastructure of practice that AF was creating is as a haven—or a place of shelter, refuge, or safety, if not a place of rest—that buffered professionals from the disorder of the broader educational system so they could engage in the intense work required of them, collect evidence of their influence on student learning, and develop a sense of personal and collective efficacy by helping them to feel they were contributing to something greater and more sustained than their singular efforts. Indeed, in response to a question about what brought them to AF, many school leaders and teachers first described the feelings of frustration or even futility they experienced in previous work settings. Take for example, the description an AF

Kindergarten teacher about why she sought work in AF after teaching in two other urban school systems: "I just felt like I always worked really hard with my class and then when they moved on to the next year , […] their teachers in first grade weren't building on that." In addition, in her previous schools she struggled to obtain the basic resources she felt she needed in her work like sets of books to engage in guided reading instruction. She explained that she and one of her co-workers went through every classroom in the school but could not put together adequate sets of books. After finding little support from her school leaders with the problem, "I then went to Reading A to Z, which is a Web site where you can print out books and I did that, but then there was no copy paper at the school to have the books photocopied, so I would go to Staples and run them off myself" (Interview33).

Similarly, an AF principal situated her narrative about what led her to AF within an account of the disheartening experience in the school in which she had previously worked as a teacher and grade level leader for several years. She recalled two specific interactions with colleagues that catalyzed her decision to leave. The first was with a teacher in her grade level who asked her, "Why do you keep doing all that, why do your lesson plans look like that, why do you have all these unit plans, why do you stay extra? […] You need to stop doing that because we're going to have to stay, too." The interviewee believed this colleague was implying, "If you put in a little bit extra then everybody has to put in extra and we don't want to do that so you need to stop doing that; stop showing people up." The second story involved "a senior level teacher" who asked her:

> "Why do you care so much?" And I [asked], "What do you mean?" And she [replied], "They're just going to end up on the corner anyway." And honestly, after that, I [decided] this is the wrong place for me, because I won't ever forget that. Because that person was largely heralded as a great teacher; when all the first graders that went into her classroom came out reading, but she didn't believe in them, she was just going through the motions and it didn't matter that they learned

how to decode words. There was no one who believed in the children. When I think about why I got into this work and why I stay in this work it's because somebody believed in me. I feel like I'm really lucky in my own educational journey, but I wanted to do that for other kids. And what I came to realize after those experiences was that may not be possible there.

As a result of these experiences and her continued dedication to traditionally underserved urban students, she began having "this vision" of being a leader in her own school:

I knew that there would be a place where teachers could talk to each other, teachers could work together, teachers would be driven to work towards common goals, teachers would hold each other accountable and not let a first grader go for a whole year with six other teachers. […] There would be some kind of sense of, we can't let this happen to our kids, which I didn't find at the [previous] school.

This vision was still quite abstract, however, because she had not yet visited a school serving a similar population of students in which these things occurred.

This changed when, through a professional connection, she spoke with an AF recruiter and was invited to visit AF's only elementary school at the time, Elm City. Like many of the school leaders and teachers I interviewed, this principal offered an eloquent description of her surprise and delight in observing in AF for the first time. Upon visiting, she reported thinking,

Oh my God, these are my kids, and this is like what it could look like if it wasn't just one classroom, if it was a whole school where the expectations were something that the teachers accepted and were working towards. That whole common vision, that whole—the language that the teachers were using, you could hear it echoed in classrooms. I saw teachers on their prep looking at their student data and I [just thought], this is amazing. […] It felt so good to know it was possible, and that you didn't have to be the only one to think that way and that you could work with other people who thought that too. And so I was just so moved on so many levels.

Finding like-minded colleagues working together with shared tools and shared aims served as an important source of inspiration and motivation for her that she did not find working in her previous placement. But since the AF environment was so distinct from the one in which she had worked before, this principal found she had a great deal to learn about how to function within it

before she was prepared to assume leadership of a school even though she was already certified to be a school administrator. She explained she needed to "learn the language" of AF, and to "understand why things happen the way they do" there, "because I'd been working for four years in a very dysfunctional place and I need to know those things before I can lead a group of people to do those things" (Interview49).

Similarly, an academic dean explained that when she first applied to work at AF after several years of teaching elsewhere, she wanted to pursue a school leadership position. Yet after visiting an AF elementary school, she reports realizing,

> My frame of reference was so different, […] I needed to reset or relearn what it means to have high expectations. I think a defining moment was when I was watching a Kindergarten class and they are reading, and I had to ask […] the recruitment director at the time three times, […] is this Kindergarten? […] I was just stunned. I mean, it was May; I was teaching second grade at the time and I could see that they were reading so much better than a lot of my second graders. (Interview31)

Ultimately, this dean taught for a year in AF before assuming the role of academic dean and strongly believed that it had been an important year of learning and growth for her.

Other school leaders and teachers provided similar rationales for leaving the traditional urban schools in which they had previously worked and highlighted the powerful resources that practicing in AF afforded them. For example, the ways that AF organized network- and school-level instructional leadership roles so that key responsibilities involved in running a school system were shared across the network in ways that protect teachers' and instructional leaders' time meant that educators could focus their efforts on work more directly related to student learning. AF's network structure allowed AFNS to take on significant responsibility for back-office functions like "teacher recruitment, fundraising, budgeting and fiscal operations, data management, information technology, and facilities operations," as well as important aspects of

teacher and leadership development (AF, n.d.). Similarly, AF's school leadership model protected instructional leaders from many of the distractions that might threaten to consume their time. A director of operations assumed responsibility for many of the schools' non-instructional functions, and a dean of students took on much of the work around student behavior, family engagement, and cultivating a strong school culture. These arrangements allowed principals more time to focus on their roles as instructional leaders, and as the coach of other instructional leaders at the school, including at least one academic dean and perhaps a principal-in-residence.

These features of the AF model seemed to feel very important to school leaders in particular. Discussing the relationship between AFNS and the network's schools, one principal commented that be believed the partnership between AFNS and the schools was strong and that they tended to function as a team, largely because of their shared goals and mission: "It is in stark contrast to the typical feeling of school versus the district that a lot of traditional schools have, where the district really is like 'the man' and […] is totally removed from the working of a school." The specific support AFNS provide to him, such as recruitment, fundraising, and support with the budget meant that his "job can really be all about the school and what happens in this building, the kids, the teachers, the parents, the instruction, the culture, and not a bunch of other administrative hassles which can take a lot of time" (Interview44).

Still others highlighted the ways in which working in AF had offered them important opportunities to learn and grow as professionals that they found to be largely absent in their earlier roles in other educational systems. Although she had previously taught for five years in another urban school district, one of the elementary school teachers I interviewed explained that her first year at AF "was almost like my first year teaching again." In her previous district, "everything was scripted" and focused on classroom management, and though she knew she had

a great deal of room to improve as a teacher, she was unclear about how to do so in that environment. She reported, "When I came here [to AF] last year it was definitely challenging, but in a completely different way. It wasn't anything about management anymore, it was much more about instruction." Though the transition was difficult, she was pleased with the opportunities she received to grow professionally: "At my old school, professional development was a joke, we really didn't get professional development, and here it's—I mean, that's all we get" (Interview34).

Another teacher working in the same elementary school described her concerns about the absence of meaningful feedback on her teaching in her original placement school in one of AF's host districts after she finished her two year stint as a TFA corps member. "My third year it was just me, you know, I'm no longer Teach For America, and [there] was just this huge decrease in my ability," she explained. She felt that in that school, she was presented with a blank slate for each day, and for the year, and was being told, "do what you will, and I didn't really know what I was doing very well." This situation had serious repercussions for her learning, and that of her students, especially in mathematics: "We used this math curriculum that I did not understand. […] And so a lot of times math would turn into, guys, […] let's just do multiplication again. I felt like at the end of the year my kids didn't really know a lot of math stuff because I just didn't teach it because I didn't get it." AF provided a very different environment: "People are coming in and out all the time in a very positive way, helping you out, offering you suggestions, you can work with people. To me it's night and day." While at first it was "nerve-wracking" to engage in such imperfect performance publicly, she reported it was not long before she learned she could say, "I don't know what I'm doing and no one's going to say, oh, my gosh, why don't you know what you're doing; they're just like, oh, let me show you. Last year a lot of times [my coach]

would come in and teach the lesson for me and say, look, this is how you're supposed to do it"
(Interview32). Her comments suggest that while AF's infrastructure may have served as a buffer
from the typical blizzard of instructional guidance that often confronts teachers in their work in
the United States' famously decentralized system (see, e.g. Cohen & Spillane, 1992), it may also
have protected teachers from the absence of guidance and feedback.

Taken together, the collective sense of purpose, dedication to continuous improvement
for both individuals and the organization, and set of resources to support such development that
staff members reported finding in AF seemed to provide many employees with an important
sense of "fit" or belonging and satisfaction in their often-challenging work. A principal's
description of the relationship between AFNS and their schools captures some of this:

> I really do think that everybody is here for the same reason and that we for the
> most part have been able to build a feeling where it's like one team. I mean, your
> job might be as principal of a school or as a third grade teacher or it might be on
> team data […] or it might be on team recruitment, but the things we have in
> common, of what brought us to AF and why we're working as hard as we work
> transcend those different roles. (Interview44)

Teachers generally reported feeling similarly. The comments below are from two teachers—the
first a first year TFA corps members, the second a more veteran teacher who had previously
worked as a teacher and school leader in a variety of school contexts:

> I was attracted to Achievement First because of the mission, because of the real
> work that was being done to close the achievement gap. I've been happy here. I've
> been very happy here as a professional. I've learned a lot [laughs]. I've learned a
> lot about myself in my professional growth. It's a place that really pushes you to
> be better than you already are. […] I don't think the school is perfect. I don't think
> any school is perfect. But I feel like I fit well here because there's always a drive
> to be better and there's an acknowledgement that what we're doing is not perfect,
> [but we're asking], how can we make this work for all of our kids? (Interview 36)

> I'm part of an organization and a school that's really working hard to make a
> difference, not just for this community, but overall, in general. And it motivates
> me and I think I'm able to then transmit my excitement and passion about what
> I'm doing to my kids. (Interview42)

These comments suggest that these educators saw the inspiration and drive that they derive from their work in AF as benefitting not just themselves, but also the students with whom they work. The creation of a community of like-minded, purpose-driven individuals seemed to offer many employees with motivation, inspiration, and a feeling that they could maximize their professional impact with ready resources for learning and improvement. It also helped to create and sustain a culture that has a reciprocal and mutually constitutive relationship with individuals' practice, in that the culture supported people in practicing in ways that reinforce this culture.

…

In sum, AF's infrastructure of practice and student achievement results often provided network leaders, school leaders, and teachers with much needed encouragement, motivation, and mechanisms for learning to do the hard work of teaching and instructional leadership, and served as a haven for those frustrated with the dysfunctions of the broader educational system that often drove them from traditional district schools serving similar student populations. In other words, AF's infrastructure and culture appeared to create space for professionals to learn and work, to maximize their effectiveness, to avoid having their strengths "trained out," and ultimately to stay in the education system longer than they might have otherwise.[90]

Lens IV: Infrastructure of Practice as an Impediment for Individual and Organizational Growth

In the previous three sections, I focused on the many affordances of AF's infrastructure of practice for supporting individual and organizational performance. As it was used within AF, this

---

[90] A senior leader at AF explained that he believed that many of the qualities AF looked for in teachers had originally been present in many more people entering the profession, but may have been "trained out" of them by working for a long time in a traditional dysfunctional setting and who essentially gave up on trying to problem solve, get feedback, or collaborate when rebuffed repeatedly by colleagues and administrators (Interview15).

infrastructure structured opportunities for individual and organizational learning, and organized the outcomes of many of those efforts; fostered processes of "mindfulness" that characterize high reliability organizations and provided resources that form a safety net to support individual and organizational performance; and sheltered professionals and their work from some of the chaos and disorder of the broader educational system in ways that motivated people despite the myriad challenges they faced in their work.

At the same time, some AF staff members—and particularly veteran school leaders and teachers—described aspects of AF's infrastructure in many of these positive ways but also had concerns about the ways that aspects of AF's infrastructure could come to act as impediments to professional sustainability and growth—their own, and that of the network as a whole. They worried about their abilities to balance personal and professional goals, and to continue to improve and deepen their practice after they had mastered the basics of their roles in AF. Some also expressed concern about some of the ways that AF's infrastructure might limit their students' opportunities to learn as well despite the network's ambitious goals for them. These observations provide a more nuanced view of AF's infrastructure of practice, culture, and the ways that they are deployed, and therefore offer insight into some of the possible challenges of developing and using an infrastructure of practice in a rapidly expanding educational system, which I discuss in more detail below.

The challenge of sustainability
AF's recruitment and selection processes yielded a staff comprised largely of novice or early career teachers who are frequently highly educated young people with little experience but a strong commitment to the organization's mission and time to dedicate to its pursuit. Network and

school leaders explained that while it would preferable to find people who exhibited both the

skills and mindsets that AF values, those candidates were few and far between; when given the

choice, the network's bet was on the importance of mindsets over skillsets in part because they

believed it to be easier to teach the former over the latter. However, despite the supports that AF

provided to their teachers to bolster the pace of their learning via their emerging infrastructure of

practice, working at AF still required an enormous amount of time and effort. As more

experienced and slightly older teachers and school leaders contemplated starting families, many

expressed trepidation about what this might mean for the sustainability of their careers in the

network, and what these aggregate losses might mean for the network.

In her response to a question about areas of growth for AF, a teacher who had been

teaching for nearly a decade shared worries typical of the teachers I interviewed who had been

teaching for three or more years:[91]

> I don't think it's a fault of [AF's], but I do think that they need to create situations
> that are more sustainable for people who are older. And when I say "older," I
> mean in their thirties. After we get married, after we have children, it should still
> be a place that's an option to work at. And I don't think that people leave because
> they're tired or they don't want to work as hard, I think they leave because […] it's
> not an option. I always say, you come in at 7:15, that's the last minute you can get
> here, and you can leave at 4:05 and that's the first minute you can leave here. If
> you do that you're not doing your job because there're so many other things you
> have to do. Besides that, I couldn't imagine at this point having a child to go home
> to and having to take care of that child. I would—something would suffer. So I
> think [there] should be more options to keep more people, because I think we lose
> a lot of good people too soon who […] know what we are and what we do and
> have learned by trial and error and have a lot of good things to say and teach
> others. […]

---

[91] Another teacher from a different elementary school expressed a very similar idea, offering that the AF workday
felt long and challenging. She continued, "I think it's a lot easier when you're 23 years old and you're single and you
don't have kids […], but I think as your teachers mature and as they get older and they get married and they start to
have kids, the structure as it is now is not sustainable in order to retain—to keep your teachers." When I asked if
these concerns factored into her own thinking about the future, she replied, "It does. I mean right now, I'm single, I
don't have any kids, and so it works for me. But I can't imagine—I talk to my friends that are married and that have
kids or going to have kids, and it's not—it just—it isn't a sustainable career, you know? Because you're here more
than you are at home sometimes" (Interview34).

SR: Is that something you're worried about for yourself in the future?

I: Yeah, absolutely. People always say, how long are you doing to stay there? […] And I always say, as long as I can, I'll be here as long as I can. But I've had so many friends who are my age who have since left for similar reasons, they've gotten married and they want to have children, or they've had children and they've had to go. (Interview43)

Describing the way that she had recently observed a principal in the network shorten a strong teacher's schedule after she returned from a maternity leave so this teacher could continue working at AF, she added that she was relieved to see this happen because of how much educators learn from each other in AF and what losing such skilled people might mean for AF.

Several school leaders shared similar questions about the continued viability of their work. When asked about the sustainability of the workload she described to me, a principal shared, "I think it's a common question you hear among teachers [and] school leaders […]. [W]e have not figured it out yet because we're just still so young, both as an organization and as a broader movement, especially with the networks of charter schools." She then referenced her personal concerns in this area:

> I don't know how much of this is just the people that are attracted to work at an AF, but I think about school all the time. My conversations with other people are about school. I take so much of that home and I'm already physically here for so many hours of the day. And then it doesn't seem compatible to me, at this point in my life, to actually move into the next phase of my life with a family, being in this position, without feeling a lot of guilt, like I'm leaving things aside and not paying attention to what I need to do, feeling like I'm not putting whatever it takes into closing the achievement gap. It's a common thing that I think that comes up in our conversations; how do we figure that out? I don't know. (Interview number redacted)

She noted that particularly when in came to teachers starting families, there was a "a lot of watching" going on in the network as AFNS and various schools experimented with ways for teachers to be able to continue to work in AF while parenting by reorganizing to use teachers' time more effectively, or by altering schedules and sometimes reducing pay accordingly. Yet

some people within the network raised questions about the extent to which creating alternative options for teachers would help with retention given the backgrounds of people who AF tended to hire—TFA or TFA-like teachers who are often people who did not enter teaching imagining it would be a career, whose families may not be local, or who were likely to be interested in pursuing advanced degrees of some kind that would take them out of the classroom.

Challenges that may accompany a strong, novice-focused system

The attrition motivated by challenges with sustainability and other turnover typical of urban school systems, together with AF's recruitment and hiring priorities, enormous growth, and practice of encouraging successful teachers to move into leadership positions after 2 or 3 years meant that approximately 42 percent AF's teachers were new to AF in 2009-2010, and 77 percent had five or fewer years of experience (Curtis, 2011). These features of AF's staffing model raised a second major set of concerns about what such an intense focus on novice and early career teachers meant for instructional quality across the network, and for the development of more experienced teachers and school leaders.

AF's leadership selection model helped to ensure that the network was able to create multiple and overlapping site-based instructional leadership positions filled by people already steeped in, and apparently successful with, AF's infrastructure. Yet several senior network leaders, and a handful of school leaders I spoke with, shared their unease that this staffing model had the potential to institutionalize a relatively shallow version of AF's vision of instructional excellence across the network. For example, if (a) after only 2-3 years of experience in the classroom, such leaders were identified and selected largely for their success with the basic features of the relatively content- and grade-neutral Essentials, and their students' proficiency on

assessments that many in the network worried were not particularly rigorous; and (b) these

emerging leaders had not yet had the opportunities do develop a deep understanding of pedagogy

or pedagogical content knowledge; and (c) the people who they led were predominantly novice

or early career teachers who needed the most support and intervention on the basic features of

the Essentials, so their leadership experience was largely centered on these basic features of

instruction yet still required enormous amounts of work.[92] If then (d) many of the more

experienced educators and school leaders moved out of their roles in classrooms or schools, or

out of the network entirely, taking much of their knowledge, experience and questions about AF

with them when they left, then it seemed possible that (e) the network risked reifying an

infrastructure that was geared toward supporting teachers to become reliably adequate in the

classroom while still working tremendously hard, but that lacked the capability to realize its

broader goals. Finally, if (f) these results were good enough, and AF's staffing system insular

enough, these performances might come to be understood as the destination rather than a first

step on a journey preventing AF's leaders from making sense of these situations in ways that

might lead them to intervene in this cycle.

Some of these concerns seemed to be reflected in coaches' discussions of their areas of

strength and areas for growth. One first-year academic dean who had previously taught for two

years outside of AF as a TFA corps member, and for a year within AF before assuming his role,

explained that he believed his areas of greatest strength was general teaching feedback, since

"good teaching is good teaching" across content areas; he offered that he was least successful in

---

[92] For example, in the Coaching the Essentials draft AF was developing during the 2009-2010 school year, a potential scope and sequence for coaching teachers was proposed that broke the Essentials into five stages of development: (a) "basic management"; (2) "core instruction," or clear aims, exit tickets with high level of mastery, factual accuracy of content, and I/we/you lesson structure, clear expectations for student verbal and written work, and opportunities for cumulative review; (3) "high engagement" strategies; (4) "rigorous instruction" (including more advanced work with content, student thinking, and differentiation); and (5) "high investment and character education." However, many of school and network leaders interviewed pointed out that in a teacher's first two or three years of teaching, many of them would be largely working on stages 1-3 of this sequence.

his work in specific content areas, and early childhood literacy in particular—though he was

working on learning as much as he could quickly, relying heavily on AFNS resources, his

principal and the school's second, more knowledgeabe and experienced academic dean to help

him to support his teachers in these areas (Interview29). Similarly, when asked if she felt equally

strong coaching teachers in different content areas, another academic dean explained,

> I don't think I'm an expert in any given subject area. Where I would say I have
> more knowledge is on things that are more removed from content. So looking at
> good teaching strategies for engagement, for management, for things like that.
> And then I have just what I've learned about guided reading, about math
> instruction, about writing, which has come from a lot of outside professional
> development, so going to a session by Lucy Calkins, going to a session by
> Fountas and Pinnell. But I wouldn't say my role is even constructed to be an
> expert in content areas.

This dean reported that she had gotten "really positive feedback" on her coaching, yet observed,

"I think a lot of that is due to just the relationships I've built with teachers. I think people feel

supported and they feel like we're working towards something." Nonetheless, she reflected,

"feedback is a funny thing because you can only give feedback on what you know, and if you

don't know of anything better, then this is kind of what you see as good, so I think there's always

room for growth there." At the end of her interview, when asked if there was anything else she

wanted to add about AF's strengths or areas for growth, this dean observed:

> I just think that as a network we're really young. […] We're all working off of
> what we know, and I mean, I'm a prime example of someone that doesn't have a
> ton of experience. I made good growth with my kids, but it doesn't mean that I
> have a lot of expertise in different subject areas. I know what it takes to get kids to
> where you want them to go because of just, the work, the effort and the looking
> for resources and that kind of "whatever it takes." […] I think that's kind of a
> question mark in it all, you know? And whoever is your coach is probably coming
> from whatever their perspective is, and there are certain things that we're trying to
> norm across the network that are not really normed. (Interview30)

Implied at the end of this statement was a "yet"; she offered Understanding by Design and

Fountas and Pinnell as two areas that she thought that coaches and then teachers across the

network should become expert, though she noted that AF was trying to leverage existing resources like these and Lemov's Taxonomy, to work towards more detailed visions of cross-network instructional excellence.

Echoing this dean's observation about her ability to produce results without a great deal of experience or expertise in content areas, one AFNS leader noted that she sometimes wondered if paradoxically, the strength of the people AF employed might in fact be an impediment to the network's development of improved supports for teacher learning and expectations about instruction in the network because they obtained solid student achievement results despite the uneven quality of teaching across the network. The leader explained:

> I think that by being a results-oriented organization with pretty ambitious, results-driven people, those [qualities] help people do whatever it takes to make sure their kids can learn, right? So I'm going to pull out kids for extra support, and I'm going to call their parents and get on them about doing their homework, and I'm going to look at my interim assessment data and figure out what went wrong and come up with a plan to address it. Even just having all of that information, results oriented people can really help push the student achievement results. That doesn't mean they don't have to be skilled, but I think sometimes it takes our teachers longer to get where they need to go, but they get there. Just because of not having all the depth of content knowledge background, or [being] totally skilled on the ten essentials. […] I think our better teachers […] are much more efficient. They still work very, very hard, but they, they're nailing the right strategy the first time so […] it's just not nearly as much. (Interview17)

She believed this state of affairs, which might be seen as an undesirable consequence of AF's infrastructure as a safety net, in turn contributed to problems of sustainability for individuals throughout the network, setting AF up for future teacher churn and the need to focus on novices.

In fact, many of the veteran teachers and leaders that I interviewed expressed concerns about their ability to continue growing and developing within a network that largely catered to novice or new teachers. For example, one teacher who had taught for several years before joining

AF, and who had been with AF for several more, was also a grade level leader and coach.

Though she reported having learned an enormous amount in her first years teaching in AF,

> [T]here is a ceiling when it comes to the training for teachers that are veterans. It's not that AF is not trying to provide it, it's that when you have the masses and most of the masses are first and second year teachers or third year teachers, it's almost like you have to do it, you have to cater to them. And I get that, [but] a lot of the professional development that I go to now on AF PD days is stuff I've seen already, more than once. There are videos that I've seen probably 15 times. And the interesting thing is you'll sit in a room and people will lean over and say what's on the screen before it happens because everyone in the room has seen it at least three times. And so I think that that is one of the growth areas that AF has, which is how do you cater to the needs of the veterans who are still trying to learn, while still being able to teach those new teachers the basics. (Interview41)[93]

At another point in the interview she noted, "If I were to give advice for professionally developing the veterans so that they feel refreshed every year and so that they want to stay, it would be to ask them what they want to learn and then to bring in outside people."[94]

The same principal quoted above wondering about the sustainability of AF's model for school leaders and teachers expressed similar sentiments about the limits of the AF infrastructure for her development over time. Discussing her appreciation and respect for AF, she commented,

> We are given a lot, and that's the power of the network: These are templates and samples, and here's curriculum and we've spent the time to research this for many, many years across the nation looking at excellent schools and these are the takeaways that we have packaged up and now you take and you do. And I tell you, I don't think that I could have done it any other way because I was so new to

---

[93] She also noted that the AF staffing model could contribute to veteran burnout. She submitted that at the beginning of the school year, when "half the people don't know what they are doing," coaching and serving as a grade level leader is "very tiring" and "exhausting." She added, "it's kind of a given that the veterans are going to own a lot [of the instructional planning] at the beginning, but then what happens is a lot of the veterans burn out very early. There's a lot of veteran burn out." Still, she noted, "the learning curve is so fast" that after the first data day cycle in November, new teachers were beginning to catch on, and after the second set of IAs six weeks later, they were able to take on increasing responsibilities within the team (Interview41).

[94] She continued, "A lot of our AF PD days are run by people that are AF-ers; […] you go to workshops and a lot of them are people that work for AF or for KIPP, which is very similar. But my advice would be bring in somebody who's a master in teaching different learning styles, you know, and invite the veterans to that, or bring in someone who's going to teach you what to do when you see a deficiency in a child, who is going to give you strategies. I mean, I signed up for workshops thinking that it was for the struggling reader and I get there and it's like, 'How can we motivate them? […] Let's give them bookmarks.' And I'm [think] no, I want to know how to teach the struggling reader, how do I get them to read, you know? […] But that would be my advice, is search the globe to find the people who can teach the veterans and bring them in.

school leadership. And to me, at the time, especially in year one and year two, that felt really supportive, to know that I could tap into this resource base from around the network and reach back out to [other AF schools] and say, 'We're coming across this challenge with Saxon [a mathematics curriculum] right now, what did you guys do when that happened?' and to share ideas in that way.

However, by the time she and her staff entered years three and four of their school's existence, she reported that some aspects of AF's infrastructure of practice began to feel less like supports and more like constraints.[95] For example, she wondered, if she decided on a different math curriculum, even though the network would allow her to do so, "would I not just have to, sort of break it up so it would fit into the scope and sequence [and] IA units?" In asking questions like this without "technical fixes," she reported, "I felt the walls around what it means to be part of a network." She added, "If something is so supportive and you're given so [many] tools, once you start making sense of that for yourself you want to have the space to create within that. And I don't know that initially I felt that. […] I feel it every single day now." Sounding almost reluctant to criticize AF, she continued, "I do think, however, that the organization is very open to feedback, and so the feedback, both that I feel like I put forward and my teachers put forward, is something that I think now is going to be acted upon." Still, AFNS' response time to these questions was slower than she wanted, perhaps, she explained, because of the amount of organizational attention that was understandably being devoted to the network's rapid growth and to the teachers and school leaders new to AF or to their roles that this growth entailed.

The question of how to balance "something so supportive" on the one hand with "the space to create within that" on the other was one that was a frequent topic of conversation in the network (see, for example, the disucssion of the Shared Practices document in Rosenberg, 2012).

---

[95] Elsewhere in the interview, this principal explained, "A parallel time frame that I think of is my third year of teaching. […] I put down the curriculum books, […] and I started noticing things about kids and grouping them in different ways and trying things out in the classroom, and I think that, at least my perception, […] doing that is harder to do in a network because there's so much that is very similar across all of the schools."

At the same time that AFNS wanted to capitalize upon the affordances of working collectively

with shared tools, they also valued and hired individuals who demonstrated unusual initiative and

skill, and sought to allow practitioners room to practice in ways that might attend to the

particular contexts of their work, to feel ownership and investment in their work, to deepen their

understanding of teaching through their experiences with learning by trial and error, and generate

ideas and practices that might benefit the network as whole. As AF evolved, grew, and recruited

large numbers of new teachers and leaders to join their ranks, the organization and its personnel

had to learn how to strike this equilibrium within the new parameters these circumstances

established. For several of the more experienced elementary school leaders and teachers I

interviewed in the spring of 2010, these things felt unblanced.

Another of AF's extraordinary strengths—their strong organizational culture—was also

mentioned by several network and school leaders as a potential liability for the organization's

growth and development. A strong organizational culture can be a strength for an organization

because it can serve as a mechanism for social control, supporting greater consistency across

individuals and improved organizational performance under certain conditions (O'Reilly &

Chatman, 1996; Sorensen, 2002). Yet some within the network wondered if AF's selection

model and strong culture could reduce the existence or visibility of divergent viewpoints

represented in the organization, threatening a key process of mindfulness in high reliability

organizations, the "reluctance to simplify interpretations" (Weick & Sutcliffe, 2001), and

perhaps serving to limit innovation or adaptation outside of these frames. For example, reflecting

on her own experience, an AFNS leader explained that she learned to teach as a TFA corps

member in a very "autonomous school setting where I just tried stuff out and if you probably

came in some days, I would have been a total mess. Sometimes I wonder, would I have been an

effective teacher, or would I have grown in the way I did, if I had been an AF teacher [with so many resources already provided]?" (Interview17). Another network leader stated that he wondered if the network's intense focus on results—while largely a positive influence in the organization—sometimes created conditions that discouraged people from experimenting with new and creative ways of working that might have eventually contributed to the organization's overall effectiveness. He continued, "I wonder if I were a teacher and I wanted to just teach something in a very different way, would I feel totally comfortable doing that if someone's going to come in and evaluate me on a certain set of teaching practices? […] If it doesn't produce results the first year around is that a problem?" (Interview16). He added that he had not seen evidence that would indicate whether or not this concern was warranted, though. Ironically, the principal quoted above as having begun to feel the limitations of the network's infrastructure worried about the opposite: that the network's strong results might prevent people from questioning their approaches. She noted that as she began to have more questions about the ways that things were unfolding at AF, "I didn't know if I should just keep my mouth shut and just keep going. Because that was working. And if something's working, why should I try to change something?" (Interview number redacted).

These questions about sustainability, longevity, the possible liabilities of "whatever it takes," and professional development opportunities for more experienced teachers raised some questions and concerns about student learning within AF because of their influence on the quality of the teachers and teaching to which they have access. There were several other ways that AF teachers and leaders suggested that AF's infrastructure of practice might act as an impediment for student learning, however. Some have already been discussed earlier in this essay or in Rosenberg (2012), such as the quality of the assessments by which teachers, leaders,

and other stakeholders evaluated their success; the ways in which state standards and IAs seemed to some to atomize learning objectives in some subject areas and grades (particularly elementary mathematics); the possible limitations of the Essentials as a guide for strong instruction because of their emphasis on a particular vision of instruction regardless of the grade, content area, learning objectives, or because they were not always accompanied by detailed attention to the ways they might be articulated differently in these different contexts; their use as a mindless checklist by some; or the absence of strong, systematic cross-network mechanisms for knowledge capture and sharing around instruction. Still, even many of these may be understood as being related to the retention of teachers and school leaders in AF, since, as one veteran teacher pointed out, "[I]f we constantly have this high turnover where we're getting first year teachers we're never going to get anywhere, you know what I mean? Because we're going to constantly be starting over" (Interview43).

Possibilities

The juxtaposition of the first three sections of this essay—in which AF's infrastructure is viewed as a offering remarkable resources through its various functions a framework and scaffold for indvidual and collective learning and practice, and as an outcome of this work; as a safety net for the organization, its staff members, and students; and as a haven for the young professionals who found their way to AF—with the previous section on the ways in which AF's infrastructure can also come to feel like an impediment to individual and organizational learning and growth begs the question: What sort of organization can AF become? One that continues to grow, not just in size but in the depth of opportunities they offer to their students and employees, or one that continues to work within the limitations of its current model?

One positive sign is that each of the concerns I identified above came from people within the network who did not want AF to fall short of its highest aspirations, and many of their reservations were shared throughout the different spheres of the network. For example, when asked about areas for growth or next steps for the organization, almost every network leader I interviewed specifically mentioned that they needed to do a better job supporting teachers across time in terms of both work/life balance and their continued professional development to improve retention and the quality of teaching and learning throughout the network. Although AF's teacher retention rate was above 80 percent in 2009-2010, Director of Evaluation and Organization Development Sarah Coon explained that she and others were committed to working to maintain those numbers, though she anticipated it would be difficult. "Many of our teachers have only been with us for a couple of years because we open new schools, we open new grades constantly so as people have been in their roles for five, six years, it'll be interesting to see if we can really keep those numbers," she explained. "We're focusing a lot on it. It's hard. Our teachers do work incredibly hard, there's absolutely no question about it." Coon's comments suggest AF's leadership took these questions seriously, but nor did they have facile answers to them.

In fact, when the interdependent topics of sustainability, longevity, and improved professional development for more skilled teachers arose in interviews across the network, it seemed as though AF was at somewhat of a crossroads with respect to these ideas: Was AF going to be a place focused on early career teachers and educators, or were they going to adapt their model to embrace at least some of their teachers across a longer career trajectory? Contemplating some of the frustrations that she and more veteran members of her staff sometimes felt with AF's existing infrastructure of practice, the sustainability of the staffing

model, and the depth of professional development offered to more experienced teachers and leaders, one school leader mused,

> I wonder if this [model] is more targeted towards that sort of zero to three-year, zero to five-year teacher who will need those kind of [intensive] resources, who will need a lot of it sort of given to them, who will have that kind of energy, be at that place in their life to do it, and therefore—I mean, and that's okay, right? But let's stop pretending that it's something that it's not, or that it's going to be something that the system is not set up to actually be or support. (Interview number redacted)

A senior AFNS staffer reported wondering something similar as she and her colleagues attempted to think through the implications of the network's rapid growth for their approach to professional development.

> I think what I need to fully understand better over time is, what's the lifetime of the teacher? I think we needed someplace to make a choice about how do we develop people. […] If the life cycle is short, then we're going to focus on certain things in professional development. I think what we're trying to do is lengthen out that life cycle of the teacher, so they feel like they're continually growing at AF. It's sort of is like the chicken and the egg thing, right? Do we have to have the great PD in order to make people feel like they're growing here and learning and developing?

Still, since it appeared the network was indeed interested in retaining more senior people and committing to their development, she thought the key questions for the network moving forward were about how to get new teachers up to speed about the basics faster and then to provide much more support for people to continue growing beyond that. She thought that the coaching model and "tightening up that initial teacher training so it's really focused" would help with the first point about teachers who are new or new to AF. Yet she felt less certain about how to pursue the next steps with teachers once they had mastered those basics:

> Next, how do we challenge them and push people on the content side to grow and develop? And I think there are two ways you can do that. One is very directive: here is content knowledge that you didn't have before that I'm going give you. But the other part is making people really responsible for their own learning and getting excited to learn other ways and recommending books for people to read

262

and having sessions that are more like: […] Here are some ways you can be more creative, and getting people to change a lesson that they may have done a very rote, basic way and presenting a new way of doing it that's not rote and basic. Just getting people to tap into that creativity, that I think also keeps people engaged and excited. So […] I think as we go to scale it's: How do we get the young folks up fast, […] And then really what is [the] length of a teacher's life span at AF? I think we want it to be long, [but] I don't know exactly how we do that. (Interview17)[96]

These comments were echoed by McCurry, who offered that as AF scaled up, he believed they needed to focus much more of their attention on critical questions like, "[H]ow do you get this solid teacher to really good, and the really good teacher to master? And what does […] professional development for really senior people [look like]? We're doing some, but not enough right now on that front." He added, "My hope is as we grow, […] we'll have more and more folks who are in that camp and that we need to provide more cohorts and more PD for them."[97]

Indeed, during the 2009-2010 school year, AFNS was attempting to take on some of these challenges around sustainability, retention, and development quite seriously. In addition to several smaller projects and their ongoing work to improve the strength of instructional coaching, there were two major initiatives in which AF was engaged that many believed held a great deal of promise for addressing some of these challenges: first, the evolution of the

---

[96] Earlier in the interview, she wondered about another scenario: "In some cases, I think [teaching is] just very hard work, so regardless [of whether or not] we figure out great PD and all of these things, people do want to leave the classroom eventually, because it's just exhausting. No matter how much PD and support you have, and that's a question I have a lot of the time. I used to [think] oh, we want to be able to [have] people see careers in teaching for 15, 20, 30 years […], but I don't know if you're doing [it] really well, if you can really keep that up. That may mean stepping out of the classroom at some point, […] but then being able to go back.

[97] Several AFNS and school leaders took pains to point out that while they had a great deal of progress yet to make in the areas of sustainability and improved development for more experienced and successful teachers, they had been attempting to attend to these things at least in small ways. For example, unlike KIPP teachers, Coon explained, "Our teachers don't teach on Saturdays, the day ends at 4:00, not 5:00. They don't have to carry cell phones [and always be available to students]." In addition, in 2009-2010, AFNS leaders added an optional "personal priorities" section to the PGP so that teachers and their coaches had a formal opportunity to discuss questions of work/life balance and priorities. Together with giving principals considerable latitude in the ways that they organized teachers' schedules and other staffing within the school, they hoped that these conversations might lead people to feel more supported, but also might provide principals with information that they could use to engage in creative problem-solving if major issues, such as having a child, threatened their ability to continue working in AF, and these could provide useful models for other school and network leaders.

curriculum and professional development team, and second, the development of a Teacher

Career Pathway (TCP).

Reconfiguration of team curriculum and professional development

In the spring of 2010, team curriculum and professional development was poised for a

restructuring to respond to evolving network needs. Previously, the team had been primarily

responsible for developing the network's scope and sequence and IAs, and with managing

network-wide professional development, including AF-wide PD day and regional content days

for middle and high schools. In addition, they served as in-house content consultants for teachers

or school leaders upon request. However, their capacity to do the latter had been limited by a

relatively small staff and the intensive demands on their time made by IA and scope and

sequence development. Yet at the time of the interviews, AFNS staff were in the process of

recruiting people for several new or reconfigured positions for the team, including a director of

K-4 ELA, 5-12 ELA, K-4 mathematics, and 5-12 mathematics, each with one to three associates

working with them, and associate positions in science, history, instructional technology.

In addition, they were in the process of renaming themselves the "teaching and learning

team" to signal a shift in their focus—away from the development of AF's scope and sequence

and IAs and towards building stronger systems for instructional sharing, learning about subject-

specific teaching, and the development of coaches more expert in particular content areas. While

these had always been network priorities, McCurry explained, "at first we had to build the

infrastructure, get the scope and sequence, get interim assessments up." One set of projects the

existing team members had already begun involved creating video libraries of effective practices

to be used in professional development and engaging in more involved studies of a handful of

exceptionally strong teachers across entire instructional units. Team members were collecting

video of their instruction, their plans, visual anchors, assessments, and examples of student work,

and then were attempting to analyze what about the instruction appeared to be driving success

and to figure out how to best share these resources and insights across the network.

A network leader explained that AFNS was working through various ideas about

developing a more "dynamic" system than their current shared server to collect and share

resources like these. She explained that she and others were imagining something in which "I

post my unit plans, I post all my lesson materials, my coach is allowed to give me feedback

directly onto that site, then I can publish them so that everyone in the network can look at them."

Still, she pointed out, "It seems really simple, but [this scheme] only works if people are

incentivized to post." Although she and others were trying to work through some of the

technological and social challenges associated with this approach before beginning this work,

she thought that having even a very underdeveloped version of this would be "better than having

nothing" and that as teachers and deans got used to using it, and AFNS staff to facilitating it,

they would all get better at what they were doing. She argued that having these resources might

be especially important for people coaching struggling teachers, but that in general, this kind of

system "is nice because then you're sharing classroom tested materials with results connected to

them, versus a unit plan I might create that might have some good components, but because it

was divorced from the classroom, it may not [prove as useful]" (Interview17).

At the same time, these plans appeared to be accompanied by a pervasive awareness that

only so much could be codified in a set of written documents and that people needed to work

with others to make sense and use of them in the particular contexts of a teacher's development

and work. School and network leaders repeatedly mentioned their convictions that handing a first

year or struggling teacher even excellent plans without thoughtful and sustained professional development and ongoing learning opportunities with more experienced others would not do much to foster the quality of work they wanted to see across the network. As Ferrell explained, "Just having a big repository of stuff, even if it's well organized, is still not necessarily sufficient in terms of adoption and usage." He continued,

> I don't think that will work, and I don't think it will work for a few reasons. […] That lesson plan might not really be what drove your outstanding student performance in that standard, it could really be […] you were just a very good teacher, you've got great classroom management, there could be elements of the lesson plan that are transferrable, but it is not the end all be all of what your instruction was for those student. So you could end up having a lot of these materials linked to these standards, but if I'm a struggling teacher I could look at all those materials and not get any benefit from them, I'm not going to move the needle in terms of my student performance on them just because I have your lesson plan.

Instead, he proposed, new or struggling teachers need access to a coach who can help them to diagnose both their most important areas of weakness as a teacher, and the reasons their students are struggling with the content, and then to provide that teacher with the support he or she needs to improve as well as helping them to select appropriate instructional materials and helping them learn to use them well. McCurry explained that in 2010, the network was in the middle of what he believed to be a three-year process to establishing the basics of coaching network-wide:

> Last year was really about, does every single person have a coach, are they meeting regularly, is the coach having some impact, do people feel good about that? And this year it's more around, […] how do we take coaching up to the next level, and can we push people to have super clear coaching plans? And I think next year is really about, alright, can almost all of our schools, and almost all of our coaches, really be at this higher level of [having a strong] coaching plan for every person, are the principal's seeing themselves [as leaders of the] coaches, are assistant spending a lot time on that?

McCurry and other AF leaders hoped that as the network more firmly established these features of their coaching model, they would be able to continue to refine the instructional coaches' role

266

and learning opportunities so they in turn would be able to provide more effective and

sophisticated support for teachers, and particularly after their first several years in the network.

The teaching and learning team was poised to take a leadership role in this respect, as

well. An academic dean with a strong focus on supporting the teachers in her portfolio in

achieving mastery in early childhood literacy practices explained her understanding of how this

vision for the team was shaping up.

> Previously, academic deans' development has been owned by the talent team, or
> the leadership development team. The academic dean trainings are all about how
> to manage people, how to have difficult conversations, how to plan an initiative;
> all great things that academic deans need, but not the nerdy content stuff. Next
> year academic deans' […] development will be owned more by the teaching and
> learning team, which is a lot more about content. And the new director of teaching
> and learning, […] this is her first whole year in the job, she's very focused on
> content and thinks that that's the next thing to infuse. So now, principals and
> deans of students will be owned more by leadership development and while
> academic deans will still get leadership development, they'll also get a big
> infusion of content specific development from the teaching and learning team and
> they'll be bucketed into grade and subject areas so that it can be more specific. So
> I think there is a network priority to infuse more content and leadership
> development, [but] it's just in its early stages. (Interview39)

This dean believed it was "very smart" that AF, as a K-12 network, placed their initial focus on a

set of Essentials "that is truly applicable to everyone," but she also thought that these plans

represented an important indication of the network's evolving priorities.

Teacher career pathway

AF was also poised to launch a new program called the Teacher Career Pathway (TCP) that

many within the network hoped would enhance their resources for supporting more experienced

teachers in their work, contribute to the sustainability of a teaching career in AF, and ultimately

to improve the quality of teaching and student achievement across the network. During the 2009-

2010 school year, AFNS leaders, school leaders, and teachers developed and piloted the program

in five schools with 30 teachers. AF (2011), Curtis (2011, pp. 15-20) and Doyle and Han (2012,

pp. 96-105) provide detailed and useful overviews of the TCP, and although it was implemented

after the year of this study, I offer a brief summary of the program here because it helps to

illustrate some of the ways that AF was grappling with these important questions about how to

measure and reward teaching excellence, and how to retain and develop successful teachers over

time during the study.

The TCP was designed to recognize and support teachers across careers within the

classroom. The TCP is organized around AF's Teaching Excellence Framework (TEF), a new

tool developed specifically for the program that aims to assess more objectively and precisely

areas of teacher performance and growth previously captured in the teacher PGPs. Curtis (2011)

explains that the TEF employs four major components to frame teachers' evaluation for the

career pathway. These include two "inputs," (a) *quality instruction*, as assessed by four annual

unannounced lesson observations using a 5-point rubric based on the Essentials, conducted by

school and network staff, and a school leader's more comprehensive assessment of teachers'

mastery of the Cycle; and (b) *core values and contributions*, assessed by a peer survey and

principal or dean's assessment of the teachers' demonstrated core values and mission related

team contribution; and two outcomes, c) *student achievement*, measured by both value-added

assessment data and principals' assessments of the accuracy of these results; and (c) *student

character development*, measured by student and parent surveys as well as lesson observations.

After a great deal of internal debate, these four categories were weighted differently to

reflect and shape the network's values. For classes in which standardized achievement data is

available, value-added scores count for 40 percent of the assessment, quality instruction and

planning 30 percent, and student character and core values for 15 percent each. In the case of

classes for which standardized data is not available, the corresponding percentages are 20 percent for student achievement, 50 percent for quality instruction and planning, and 15 percent for both student character and core values. In addition to a holistic cut-score based on these distributions, each category also has a minimum required score that teachers must achieve to move across a five-stage career path. The career path begins with Stage 1, "Intern" for the small number of people who are hired into this new role in the network. Most new hires begin at Stage 2, "Teacher" and "solid contributor." After two years of successful results, at the principal's discretion, teachers can advance to Stage 3, at which point they are recognized as a "strong, stable contributor." After at least two years of meeting particular goals relative to the TEF, a teacher may move to Stage 4, "Senior Teacher," and then after at least another two years and meeting additional TEF goals, to Stage 5, Master Teacher.

Recognition, compensation, and professional growth opportunities expand as teachers move up each stage. In addition to the standard resources provided to each AF teacher, Curtis (2011) noted that teachers in stages 4 and 5 receive "a self-directed professional development budget, a senior or master teacher cohort, opportunities to visit excellent teachers regionally and nationally, and preferred access to special professional development series." They are also afforded "benefits related to sustainability [that] are aimed at making AF schools attractive to high performers. At stages 4 and 5 these include hosting a teacher intern, the possibility of course-load reduction, and maintaining the same classes/grade structure over multiple years" (p. 18). Advancement across each stage can earn teachers at least a $10,000 salary increase. School-wide bonuses based on the AF School Report Card can result in up to 10 percent salary increases for the entire staff.

During the 2010-2011 school year the TCP was piloted much more broadly across the network and was modified somewhat for the next year. The 2011-2012 school year was the first full year of the program, and data from this year will be used to place teachers in one of the five stages for the first time in the fall of 2012. As AF deploys this TCP, they continue to collect feedback from teachers, school leaders, and network leaders about its strengths and weaknesses, and to make small adjustments as necessary. In addition, they are actively engaged in the development of additional assessments that can be used across the network in areas in which they have not traditionally been available so that they can be used in more standardized ways in the TCP process and to monitor student learning. Doyle and Han (2012) note that although the salary increases associated with the TCP will initially be paid out of a 6.3 million dollar Teacher Incentive Fund (TIF) grant they received to develop this model, AF is committed to finding ways to manage their budget so that this program is sustainable over time.

Depending on how the details are enacted, the TCP appears to hold a great deal of potential for supporting teacher quality over time. First, the TCP process further articulates AF's standards for excellence in instruction and expectations for professional practice and growth. The observation rubric created as part of the TCP process provides much more detailed information to AF's members about the network's vision of instructional excellence than the Essentials or the Cycle alone, and provides clear opportunities for norming around these standards several times a year. In addition, the TCP coordinates and aligns several sources of feedback on teachers' work in more detailed ways than in the past. These sources of feedback include student achievement data, lesson observation and feedback, and feedback from colleagues, students, and families, providing teachers with a more holistic assessment of their professional performance in addition to the focused work they do much more frequently with a coach. By provide a formal career path

for people who want to remain in the classroom rather than moving into school or network leadership positions, AF is also able to signal the importance they place on teaching and to formally recognize, reward, and prioritize the development of master teachers, and in doing so, to hopefully retain and leverage some of the benefits of a more stable and experienced staff. Finally, the TCP process may serve as an important source of data for the network about where their teachers' performance falls based on these values. One network leader, for example, mentioned that through the TCP pilot process, she and others had been surprised to learn how much support most of the teachers in the network still needed to be considered proficient, which helped to strengthen their resolve to provide these teachers and their students with the supports they needed to improve and improve quickly.

Conclusion

I have argued that AF's infrastructure of practice, though still new, offers a rich model of organizing for instructional quality with several important lessons for reformers who aim to improve teaching and learning across a school system. First, AF's infrastructure provided its leaders and educators with a set of key instruments through which they could influence essential features of instruction, and with a framework for organizing these tools as well as sustained and overlapping supports for their use and improvement. In other words, the infrastructure may be understood as dynamic if still-emergent repository of knowledge (and knowing) for the practices of teaching, school leadership, and school system leadership.

Second, in part by treating student achievement as their version of the life-or-death outcomes faced by more typical high-reliability organizations, AF's infrastructure fostered the processes of mindfulness that helped to motivate the organization's continuous improvement efforts. These processes and other aspects of the organization's culture helped AF to recognize,

learn from, overcome, or contain underperformance across the network and created a sort of "safety net" for organizational, professional, and most importantly, student performance.

Third, AF's infrastructure was described as forming a professional haven for many of its employees, protecting their time, work, and beliefs from some of the frequent disturbances of their environment. This haven also provided employees with likeminded, dedicated colleagues who were teaching in similar contexts with many of the same tools, allowing teachers to distribute their workload and expertise. In all of these ways, AF's infrastructure helped to support a notion of teaching and instructional leadership as "shared, systematic practice" (Raudenbush, 2009), in which these ways of functioning were broadly seen as the default way to work and to improve, or as simply part of the work of teaching.

However, I discovered that several of the infrastructure's greatest strengths also created corresponding tensions for AF. In part because few examples of strong, cross-school infrastructures for practice exist in their operating environment, AF had to identify and build many of what they had come to see as critical pieces of their infrastructure for themselves and to adjust or add to these as they grew, learned from experience, and developed new hypotheses about what might be necessary to pursue their goals. Therefore, the infrastructure was still quite uneven across grade levels and subject areas, and, some within the network contended, overly focused on the development of new or early career practitioners—with several ramifications, including the risk that this situation could become self-perpetuating and inhibit the network's continued development. If AF continually lost the more experienced teachers and leaders who might motivate and contribute to the network's development of models that supported instructional expertise beyond the basics as they were outlined by the Essentials, the network

appeared to run a greater risk of becoming complacent with their existing (and impressive) successes but ultimately fall short of meeting their broader goals for students.

The awareness of these challenges that permeated the network, together with the existing strengths of the infrastructure and AF's sincere dedication to their students' success, provide real reasons for optimism about the network's potential evolution. So do the steps they were beginning to take through the refinement of their coaching model, reconfiguration of team teaching and learning, and the new teacher career pathway. While AF still has an enormous amount to learn, we have a great deal to learn from what they have done already.

References

Achievement First. (2011). AF Teacher Career Pathways Executive Summary. Retrieved May, 2012, from: http://conncan.org/learn/research/teachers/measuring-teacher-effectiveness

Achievement First. (n.d.). Our Mission and Vision, and The Achievement First Approach: As a Network. Retrieved June, 2009 from www.achievementfirst.org

Berends, H., Boersma, K., & Weggeman, M. (2003). The structuration of organizational learning. *Human Relations. 56*(9), 1035-1056.

Brown, J., & Duguid, P. (1991). Organizational learning and communities of practice: Toward a unified view of working, learning, and innovation. *Organizational Science, 2*(1): 40-57.

Bryk, A.S. (2009). Support a science of performance improvement. *Phi Delta Kappan, 90*(8), pp. 597-600.

Cook, S.D.N. & Brown, J. (1999). Bridging epistemologies: the generative dance between organizational knowledge and organizational knowing. *Organization Science, 10*(4), 381-400.

Cook, S.D.N. & Yanow, D. (1993). Culture and organizational learning. *Journal of Management Inquiry, 2*(4), 373–90.

Cohen, D.K. (2011). *Teaching: Practice and its Predicaments*. Cambridge, MA: Harvard University Press.

Cohen, D.K. and Spillane, J. (1992). Policy and practice: The relations between governance and instruction. In G. Grant (Ed.), *Review of Research in Education, 18*, 3-49.

Curtis, R. (2011). Achievement First: Developing a teacher performance management system that recognizes excellence. Washington, DC: The Aspen Institute. Retrieved March, 2011 from http://www.aspeninstitute.org

Doyle, D., & Han, J.G. (2012). Measuring teacher effectiveness: A look "under the hood" of teacher evaluation in 10 sites. New York: 50CAN; New Haven, CT: ConnCAN; and Chapel Hill, NC: Public Impact. Retrieved June, 2012 from http://www.conncan.org/learn/research/teachers/measuring-teacher-effectiveness

Hiebert, J., Gallimore, R., & Stigler, J. W. (2002). A knowledge base for the teaching profession: What would it look like and how can we get one*? Educational Researcher, 31*(3), 3–15.

Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge, UK: Cambridge University Press.

Lemov, D. (2010). *Teach like a champion: 49 Techniques that put students on the path to college.* San Francisco, CA: Jossey-Bass.

O'Reilly, C., & Chatman, J. (1996). Culture as social control: Corporations, cults and commitment. In B. Staw & L. Cummings (Eds*.). Research in Organizational Behavior*: Vol. 18, Greenwich, CT: JAI Press, 157-200.

Orlikowski, W.J. (2002). Knowing in practice: Enacting a collective capability in distributed organizing. *Organization Science, 10*, 249–273.

Orlikowski, W.K. (2006). Material knowing: The scaffolding of human knowledgeability. European Journal of Information Systems. (15) 460-466.

Elmore, R. (2004). *School Reform from the Inside Out: Policy, Practice, and Performance*. Cambridge, MA: Harvard Education Press.

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

Raudenbush, S. (2009). The Brown legacy and the O'Connor challenge: Transforming schools in the image of children's potential. *Educational Researcher 38*(3):169-180.

Rosenberg, S. (2012). Individual and Systemic Approaches to Teaching Quality: Building Professional Capability in a Charter School Network. Unpublished manuscript.

Sørensen. J.B. Sørensen. (2002). The Strength of Corporate Culture and the Reliability of Firm Performance. *Administrative Science Quarterly* 47: 70-91.

Weick, K. & McDaniel, R., (1989). How professional organizations work: Implications for school organization and management. In T.J. Sergiovanni & J.H. Moore (Eds.), *Schooling for Tomorrow: Directing Reforms to Issues that Count* (pp. 330-355). Boston, MA: Allyn and Bacon.

Weick, K.E., & Sutcliffe, K.M. (2001). Managing the Unexpected: Assuring High Performance in an Age of Complexity. University of Michigan Pressing Problem Series. San Francisco: Jossey-Bass.

Weick, K., Sutcliffe, K.M., and Obstfeld, D. (2008). Organizing for high reliability: Processes of collective mindfulness. In A. Boin (Ed.), *Crisis Management: Volume III*, pp. 31-66, Thousand Oaks, CA: Sage. (Reprinted from R.S. Sutton and B.M. Staw (eds), *Research in Organizational Behavior*, *Volume I*. Stanford: Jai Press, 1999, pp. 81–123.)

CHAPTER 5

CONCLUSION

The three essays included in this dissertation are intended to contribute to deliberations about what it might take to improve the quality and reliability of instruction in U.S. schools, with a particular focus on those poor and minority students who are most dependent on the opportunities they have to learn in school for their academic and economic success.

In the first essay (Chapter Two), I sought to understand why over a century of intense interest in and research about what makes a "good teacher" has failed to cohere as a more consistent and actionable knowledge base for practice and practice improvement. I found researchers have conceptualized of and studied teacher and later teaching quality in several very different ways. Teacher quality was variously assumed to inhere in (a) the characteristics or qualifications of the teacher; (b) teacher behaviors; (c) social interaction, knowledge, and cognitive processes; (d) the organizations or systems of which teachers are a part; or (e) simple measures of gains in students' achievement. While these studies and the ways in which their results were interpreted have contributed to our general understanding of instruction with some hints about several of the necessary if insufficient conditions for how it might be improved, I argued that two essential features of the U.S. educational context—including the absence of a shared set of goals and technologies for practice, and the continued influence of the political system that thwarted their development to begin with—have undermined both the usefulness of individual studies and the coordination among them that would be required to develop and sustain a widely shared professional knowledge base for teaching.

With no governmental or professional organization to establish a shared set of aims for educational systems, or the exams or curricular frames that would further articulate what might need to be learned, it has been difficult to build other features of what Cohen (2011) calls an "infrastructure of practice," including standards or norms of practice; teacher education and professional development devoted to helping teachers to learn to practice well; and a specialized vocabulary for discussing and assessing teacher and student work and its improvement. The lack of consistency across these features of the educational system means that it has been difficult to conduct meaningful studies or to build a coherent and usable professional knowledge base from their findings. For example, what does it mean to assess the magnitude of the mean influence of teacher education on student learning when "teacher education" means so many things, or when the tests used to assess student learning do not measure the same constructs as those employed in a neighboring state, or those that many believe to be essential for students' future successes? How can researchers isolate the influence of particular teaching strategies on student learning if each of the teachers in the study are required by their schools or districts to use curricula that are supposedly aligned with state standards but are actually wildly divergent from the standards and each other, with extraordinarily different supports for their use, using norm-referenced tests specifically designed *not* to be sensitive to variations in teachers' instruction to estimate teachers' "value-added"? How can they effectively coordinate and build upon each other's work in these circumstances?

Indeed, teacher and teaching quality research and policy seems to be subject to a bizarre Catch-22. On the one hand, policymakers, practitioners, and the public clamor for answers to these questions derived from science. Yet without common aims and other key resources, researchers have had a difficult time conducting research that would elucidate the practices in

277

which strong teachers engage to support deep student learning as well as the mechanisms through which teachers learn and might be encouraged to practice in these ways with diverse learners in particular contexts over time. Scholars have also encountered challenges with respect to framing questions, representing their findings, and organizing their efforts in ways that increase the chance that they might be useful in practice (Hiebert, Gallimore, & Stigler, 2002). On the other hand, without research "proving" that good teachers share particular characteristics or practices across settings, educators or reformers have a difficult time designing and getting support for the dissemination and use of the resources around which schools and school systems might be organized to improve teaching.

The second and third essays in this dissertation (Chapters Three and Four) draw upon a 2009-2010 case study of Achievement First (AF), a high-performing K-12 charter management organization enrolling predominantly poor and minority students in Connecticut and New York. Because of the critical influence of instruction on student learning and their strong commitment to providing all of their students with the education necessary to succeed in college and beyond, I investigated the ways AF worked to define, develop, and manage teaching quality across their quickly expanding network of schools. I found that AF had seized upon the opening in the U.S.' decentralized system created by legislation enabling charters and used its schools' relative autonomy to begin building a coherent educational subsystem at the margins of the traditional one as they organized to attempt to provide their students with better and more reliable opportunities to learn.

Within this somewhat autonomous subsystem, AF had several resources at their disposal that helped them to address some of the problems that have plagued research on teacher and teaching quality across the past century. For example, in addition to being sheltered from some

of the state and district mandates that have threated similar reform efforts in the past, AF also had the advantage of operating as a network of schools of choice. This meant that students and their families, teachers, and school leaders could opt to participate in this system or not since they could theoretically "choose" to attend or work in other schools. This provided AF's leaders with the freedom to unapologetically establish a set of common aims for their work—most broadly, that their schools were geared toward preparing students to succeed in college and beyond, but also more proximal achievement targets usually measured by standardized exams. AF's leadership could also draw upon prior research and their experience to select, develop, or refine a set of tools and practices that reflected their hypotheses about how to meet these targets; to establish expectations for their use across the network; and to hire people who subscribed to their approach. Having the freedom to begin with these informed hypotheses allowed them to try them out in practice, to collect and reflect upon data about how they were working, and finally to refine them, add to them, or adopt a different approach. In addition, since AF's schools were all "fresh-start" charters, they were able to build this culture and other resources from the beginning rather than engaging in the challenging work of "turning around" existing schools.

AF's professional support and management system, and the ways that the network was organizing within and around it at the time of this case study, offers an vivid example of a still-developing infrastructure of practice and some of its profound affordances for professional learning and practice. As it was used in AF, the infrastructure provided teachers and leaders with guidance about the aims, content, method, and organization of teachers' work; standards for discussing and assessing practice; and a means for organizing learning about teaching and instructional leadership. It also supported and was supported by a growth- and performance-oriented culture characterized by a great deal of collaborative work, mutual responsibility, and

trust. This suggests that the notion of an "infrastructure" provides a useful heuristic for studying and building the knowledge and skill of individual teachers and their collective capability, and therefore for supporting reliability in teaching quality at a scale rarely found in traditional school systems in the U.S.

In fact, AF's infrastructure of practice appears to have successfully focused staff in all spheres of the organization on producing and using knowledge about practice to improve their individual and collective performance, and can be understood as evidence that AF was in the very early stages of building a professional knowledge base for teaching—something that is extraordinarily rare in this country despite a great deal of interest and activity that has aimed to contribute to its construction. For example, Hiebert, Gallimore, & Stigler (2002) observe that in an effort to produce "objective" and broadly generalizable knowledge across the myriad contexts in which researchers work, much of the research on teacher and teaching quality of the kind profiled in Chapter Two generates propositional knowledge that is "represented in forms that are relatively abstract, ignore contextual influences, and isolate aspects of practice that cannot easily be reintegrated with interacting features of classrooms" (p. 11). Conversely, practitioners often engage in more or less formal investigations on their own or in small professional learning communities with the aim of improving practice. The practitioner knowledge they develop has the strengths of being linked with practice; detailed, concrete, and specific; and integrated according to the problems of practice the knowledge is meant to address. However, "no infrastructure encourages, or even enables, them to record, share, and accumulate the knowledge they construct" (p. 12). As Lampert, Boerst, & Graziani (2011) note, even detailed and thoughtful efforts to "unpack" the complex efforts of skilled individual teachers engaging in

"ambitious" teaching practice for study and learning may reinforce notions of skilled teaching as "both herculean and idiosyncratic." They continue,

> The organizational context that supports these teachers is either largely invisible or portrayed as erecting barriers that need to be overcome, making it difficult to understand how to support the broad replication of the instructional models they represent. Instead, their work appears as a peculiar piece in a professional patchwork in which each teacher constructs his or her own knowledge of teaching, uses an eclectic blend of materials, and works as an individual. (p. 1394-1395)

Hiebert et al. argue that for practitioner knowledge to become professional knowledge, it must meet three conditions. It must be *public*, in the sense that it must be "created with the *intent* of public examination, with the goal of making it shareable among teachers, open for discussion, verification, and refutation or modification" [emphasis in original]. It must be *storable and shareable* so that it may be accumulated and distributed, as with case literature in medicine or case law, and to this end should be "represented through *theories with examples*" that "offer abstract knowledge that transcend particular classrooms and contexts and ensure that the knowledge rises above idiosyncratic technique" as well as the examples that "keep the theories grounded in practice and reveal the meaning of verbal propositions" (p. 7) [emphasis in original].[98] They envision "large digital libraries linking video examples of teaching, images of students' work, and commentary by teachers and researchers, all integrated around shared topics, and even shared lessons […] linked to specific curricula a teacher is responsible to teach" (p. 8). Finally, since "there is no guarantee that the knowledge generated at local sites is correct or even useful" such resources would need to be *verified and improved* by drawing upon available

---

[98] Hiebert et al. (2002) suggested, "[U]seful theories, in this context, are teachers' hypotheses or predictions regarding the relationships between classroom practices and student learning, along with explanations for observed connections. Why was this instructional activity created to support this kind of learning? In what way was students' thinking expected to change over the course of the lesson, and why did such change (not) occur? These hypotheses or rationales begin transforming knowledge gained in one classroom into a form that can help other teachers think about how this practice might work in their contexts. Local hypotheses gradually develop into theories that can be tested and refined across a range of contexts" (p. 7).

expertise and the "continual evaluation of practices as they are shared among teachers and tested out in different local contexts" (p. 9). More traditional researchers could then build upon these discoveries to test some of these hypotheses more broadly (p. 13).

AF's focus on making practice public to support practitioner learning; the shared tools and practices available through the infrastructure that frame common problems of practice and ways of addressing them; and the network's organization as a set of overlapping professional networks tied together by mutual accountability, responsibility, and trust all provide a powerful foundation for the organization and staff to engage in the kind of professional knowledge development that Hiebert et al. describe. Although AF was still in the early stages of developing systems that would make such knowledge more widely shareable across the network and would encourage a more detailed articulation of the theories and examples through which it might be represented and indexed, the systems envisioned by Hiebert et al. to support this storage and sharing are remarkably similar to the improved "knowledge capture and sharing" systems around instruction that the network was planning through team teaching and learning. The network's proclivity for continuous improvement and the many practices and norms they had established about data collection and analysis suggested they were also well positioned to focus their efforts around validating these resources in more focused ways over time. Beyond simply developing a knowledge base for teaching, though, AF was simultaneously engaged in the work of developing a knowledge base for instructional and school leadership and system-wide leadership, not to mention the more specific work of particular teams or groups within the network.

As innovative, significant, and even inspiring as many of AF's efforts in these areas were in part because of their schools' autonomy as charters, several features of the broader educational environment also had serious implications for AF's work. To carve out a niche for themselves

and establish their legitimacy in the field—ensuring their continued authorization, the investment of private foundations and philanthropists, and their ability to substantiate their claims that if schools provide poor and minority students with adequate opportunities to learn, they can perform as well as their more affluent peers—AF organized in ways that accommodated several essential features of this context roughly as they were. These included the voluminous state content standards in New York and Connecticut, state standardized assessments and proficiency benchmarks, and a teaching force that generally included underprepared teachers who leave urban schools or teaching altogether at terrific rates. Paradoxically, it seemed that the strengths of the systems AF was in the process of designing to make schools work much better for their students within this landscape also created another set of tensions for the network to manage that some network leaders, school leaders, and teachers worried might ultimately compromise AF's ability realize their most ambitious goals for their students and for education reform.

First, as I observed in Chapter 3, the Connecticut and New York standards and standardized assessments were central to the ways in which success in teaching and learning was defined within AF. Therefore, these externally designed tools exercised a powerful backward effect on the design of the rest of their infrastructure, including the AF scope and sequence, interim assessments, the Essentials of Effective Instruction, and teacher evaluations—with what was determined to be "effective" being judged largely on the extent to which it appeared to contribute to students' performances on these tests. As they were used within AF, these tools and practices appeared to exert a strong influence instruction and instructional leadership, reflecting the network's willingness to assume that these tools inherited from the states offered worthwhile (or at least benign) supports for and feedback on their progress toward their ultimate goals for students.

Second, as I suggested in the third essay (Chapter 4), the intensive scaffolds AF was building and deploying to support their teachers, who were overwhelmingly in their first through fifth years of teaching, in attaining at least a minimum standard of practice represent enormous accomplishments and offer an important model for those who seek to enable and sustain more reliably strong teaching across a school system. Yet at the same time, AF's focus on its new teachers, while sensible, also threatened to create a situation in which the infrastructure *only* served early career teachers and school leaders by making demands on their time that were not sustainable, by failing to provide adequate opportunities for more experienced professionals to continue to learn and grow, or by frustrating the highly educated, ambitious critical thinkers they prized in their hiring process if they came to question some of the ways in which AF's infrastructure may serve to limit their practice and students' learning even if they believed its net influence for students and education reform was positive.

These two challenges interacted to create a third. In part because of this concentrated focus on novice and early career teachers, and because of the absence of rigorous assessments of student learning or standards that provide detailed guidance for teaching from the broader educational environment, the aspects of AF's infrastructure of practice that seemed to be the least developed so far were those that were most directly related to detailed depictions of the deepest parts of teaching—or the explanations of who is supposed to be doing what, when, with content and diverse students to result in student learning, and *why* (Hierbert et al., 2002; Lampert & Graziani, 2009). What is more, AF seemed to be at risk of falling into a cycle in which the network would consistently lose the people who were best positioned to build such resources: as more experienced teachers and school leaders leave AF for the reasons mentioned above or others, or leave the classroom to assume leadership positions as the network continued to rapidly

expand, more new or early career educators would have to be hired to take their places. As a result, the network was likely to need to continue directing a large proportion of their resources toward supports and practices geared toward new teachers—because there were so many of them, and because the liabilities of their failures so high—beginning the cycle again and threatening to institutionalize the loss of the people poised to develop the experience and expertise that some people reported the network needed to improve the quality of the learning opportunities they extended to their staff and students.

Over the next several years, it seems likely that AF will continue to struggle with two competing tendencies inherent in their work and the environment in which they operate. The first is to aggressively pursue their highest aspirations for their work and for their students by pushing themselves to substantively elaborate their infrastructure and to support their teachers in more deeply engaging their students in the ambitious intellectual work that would allow them to flourish and meet the network's high expectations for them. The second tendency is to succumb to the many incentives in their environment to establish their legitimacy by (a) celebrating their impressive results relative to the traditional school systems their students might have otherwise attended, solidifying their public image, revenue streams, and position in the education reform movement, and (b) scaling up, rather than concentrating their resources on deepening and strengthening their existing program.

Indeed, these competing tendencies reflect a deep adaptive challenge for AF, and as such have no simple solutions. Heifetz, Kania, and Kramer (2004) explain that in contrast to technical problems, or those that are "well defined: Their solutions are known and those with adequate expertise and organizational capacity can solve them," (p. 24), adaptive challenges are complex, "not so well defined, the answers are not known in advance, and many different stakeholders are

involved, each with their own perspectives." They contend that adaptive problems "grow out of conflicting values among stakeholders, or internal contradictions between the values they stand for and the realities they face," and continue:

> Adaptive work, therefore, requires a change in values, beliefs, or behavior on the part of those with an interest in the problem, and such changes cannot be externally imposed. The core of adaptive work is mediating these conflicts and internal contradictions, and providing the leverage that motivates people to learn new ways of thinking. Therefore, the central task of adaptive leadership is mobilizing people to clarify what matters most, in what balance, and with which trade-offs. People and institutions that lead must harness, manage, and ultimately defuse conflict among interested parties so that each can adapt to the other and to the situation in a manner that brings about social progress. (p. 25)

In all of these ways, the AF case study calls into question the wisdom of aggressively pursuing education reform strategies that are based on the assumption that high-performing charter school networks like AF will be able to scale at the pace required to substantially reduce achievement gaps in geographies across the country—because of the enormous amount of learning in which individuals and the organizations have yet to do, and because of heavy relational work they require. AF's infrastructure of practice allowed them to continue to obtain impressive student outcomes fairly reliably even as they pressed their expertise horizon quickly so they could reach more students (and, according to some, appease funders), but they were already encountering some of the growing pains associated with how aggressively they were pushing themselves in these areas and their ability to improve and sustain their work.

Despite the difficulty inherent these challenges, AF's infrastructure and culture offered useful resources for engaging in this adaptive work. For example:

- AF's *deep commitments* to providing its students with the academic and character skills to succeed in college and beyond provided some balance to their laser-like focus on more proximal measures student performance, and seemed to give many people within the network reasons to reflect upon the extent to which their educational program was supporting the full range of aspirations they held for their students even if they appeared

to be succeeding on standardized tests. At all levels of the organization, I found people who were not content to rest on their existing successes.[99]

- Though achievement on standardized tests was the primary indicator that AF used to gauge its success, this intensive focus was also *balanced* somewhat by other important *metrics*. For example, AF's "balanced scorecard," AF report cards for each school, and the new Teacher Career Path for assessing teachers' performance each included a range of measures for assessing performance.

- *Strong working relationships* formed through *overlapping professional networks* across AF, and especially coaching dyads and elementary school grade level or content area teams, provided practitioners with access to others who might help them to use elements of AF's infrastructure in balanced and increasingly ambitious ways while also distributing their demanding workload, contributing the sustainability of their work.

- As many school leaders and teachers pointed out, the AF model included a longer school day and year, meaning that educators had *more time* with students to prepare them in ways that may be easily measured by achievement tests, and in ways that may not.

- AF also cultivated the *processes of mindfulness* that typify effective high reliability organizations, and these processes did seem to be support AF's leadership in noticing some of the network's budding challenges so they could move to address them.

- Finally, despite all of their internal resources and strengths, AF was also very *willing to learn from others*, borrowing resources or practices—like the Fountas and Pinnell Continuum, or parts of Lemov's Taxonomy—and thinking carefully about how to apply them in the context of their own infrastructure and culture.

These qualities suggest that the educators and leaders working within the network may indeed be able to confront some of the tensions inherent in their existing approach. In fact, at the time of the study, AF was drawing upon several of these resources to design two important new initiatives that they hoped would more directly address problems of sustainability and development for more experienced teachers and instructional leaders: (a) the transformation of "team curriculum and professional development" into "team teaching and learning," which would focus more on developing content expertise across the network but especially with

---

[99] For example, several people mentioned the ways that the co-CEOs Doug McCurry and Dacia Toll had begun pushing people within the network to think about the ways in which their students' test scores were strong, but how, if you compared them with some of the best districts in New York and Connecticut, AF's students had many more 3's (the cutoff for proficiency) than 4's (the designation for mastery): McCurry and Toll wanted to engage people throughout the network in thinking about what it would take to push their students to the next levels of achievement.
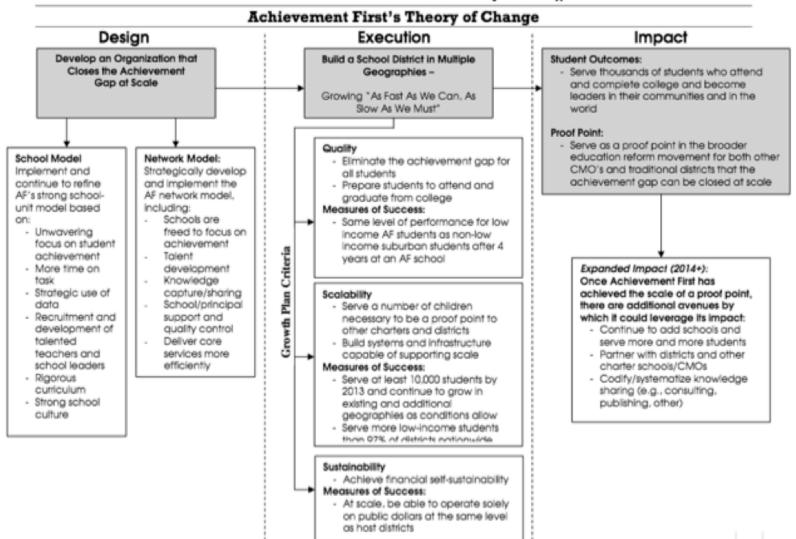
academic deans, and on strengthening the network's instructional resources and the technologies through which they were represented and shared network-wide; and (b) the development of the Teacher Career Pathway to more clearly recognize, incentivize, and articulate their vision of good teaching across the network. By extending the "lifecycle" of some of their best teachers and leaders, AF aimed to better serve their students and to create opportunities for the network to learn from their work.

The influence of another major current reform initiative—the advent of the Common Core standards and the corresponding assessments—may intersect with these rival tendencies in interesting ways as well. If the assessments of the new standards prove to be as fundamentally different from previous assessments as their designers claim and do a better job of assessing more complex skills and understandings, it may be that their implementation in the 2014-15 school year could work to bring these two tendencies together. In other words, more rigorous assessments might be productively disruptive in that they reveal whether or not AF's infrastructure is adequately preparing students for success in this new framework. If not, then the institutional incentives to perform on the exams and assert their legitimacy *might* correspond more closely with AF's ambitions for their students and provide AF with feedback that helps them to capitalize upon their existing strengths to further elaborate their program. Regardless, I am confident that AF's response will be worth watching.

References

Cohen, D.K. (2011). *Teaching: Practice and its Predicaments*. Cambridge, MA: Harvard University Press.

Heifetz, R.A., Kania, J.V., & Kramer, M.R. (2004). Leading Boldly. *Stanford Social Innovation Review*, Winter 2004.

Hiebert, J., Gallimore, R., & Stigler, J. W. (2002). A knowledge base for the teaching profession: What would it look like and how can we get one*? Educational Researcher, 31*(3), 3–15.

Lampert, Lampert, M. & Graziani, F. (2009). Instructional activities as a tool for teachers' and teacher educators' learning in and for practice. *Elementary School Journal, 109(5),* 491-509.

Lampert, M., Boerst, T.A., & Graziani, F. (2011). Organizational Resources in the Service of School-Wide Ambitious Teaching Practice. *Teachers College Record, 113*(7), 1361-1400.

Achievement First's Theory of Change

Design	Execution	Impact

Design

Develop an Organization that Closes the Achievement Gap at Scale

School Model
Implement and continue to refine AF's strong school-unit model based on:
- Unwavering focus on student achievement
- More time on task
- Strategic use of data
- Recruitment and development of talented teachers and school leaders
- Rigorous curriculum
- Strong school culture

Network Model:
Strategically develop and implement the AF network model, including:
- Schools are freed to focus on achievement
- Talent development
- Knowledge capture/sharing
- School/principal support and quality control
- Deliver core services more efficiently

Execution

Build a School District in Multiple Geographies –

Growing "As Fast As We Can, As Slow As We Must"

Growth Plan Criteria

Quality
- Eliminate the achievement gap for all students
- Prepare students to attend and graduate from college

Measures of Success:
- Same level of performance for low income AF students as non-low income suburban students after 4 years at an AF school

Scalability
- Serve a number of children necessary to be a proof point to other charters and districts
- Build systems and infrastructure capable of supporting scale

Measures of Success:
- Serve at least 10,000 students by 2013 and continue to grow in existing and additional geographies as conditions allow
- Serve more low-income students than 97% of districts nationwide

Sustainability
- Achieve financial self-sustainability

Measures of Success:
- At scale, be able to operate solely on public dollars at the same level as host districts

Impact

Student Outcomes:
- Serve thousands of students who attend and complete college and become leaders in their communities and in the world

Proof Point:
- Serve as a proof point in the broader education reform movement for both other CMO's and traditional districts that the achievement gap can be closed at scale

Expanded Impact (2014+):
Once Achievement First has achieved the scale of a proof point, there are additional avenues by which it could leverage its impact:
- Continue to add schools and serve more and more students
- Partner with districts and other charter schools/CMOs
- Codify/systematize knowledge sharing (e.g., consulting, publishing, other)

290

APPENDIX B: AF Strategy Map and Balanced Scorecard Metrics

Mission

To deliver on the often-denied promise of equal educational opportunity for America's urban children by inspiring in them the belief that they can achieve and by developing in them the academic skills and strength of character that will enable them to graduate from college and be leaders in a competitive world.

Social Impact

Direct Impact
Create high performing charter schools that close the achievement gap and enable students to graduate college and be successful in a competitive world

Metrics:
1. AF low income student achievement results vs. state non-low income student achievement result
2. High school graduation %
3. College acceptance %
4. College matriculation and graduation rates

Systemic Impact
Create a proof point of the ability to close the achievement gap at scale (size of a medium sized public district) using tools and systems which are replicable and sharable

Metrics:
5. Scale (# of students, # of academies)
6. Market share in host districts
7. Partnerships

Constituents (and value proposition)

Internal:

Students/parents
(achieve academic excellence and strength of character necessary for success in college)

Metric:
8. Student wait lists
9. Parent feedback

Teaching/leadership talent
(AF first choice for top talent who fit our target profile)

Metric:
10. # of quality applicants/position

External:

Philanthropists
(see AF as place to have max. impact with dollars and time)

Metric:
11. # of major donors

Boards
(high satisfaction with AF and assist in accomplishing mission)

Metric:
12. Board satisfaction

State and local policy makers
(provide public resources/support)

Metric:
13. Per pupil funding levels
14. Facilities provided/ funded
15. Expansion opportunities provided

Educational Model

School Model
Develop and continually refine model of outstanding educational program

Metric:
16. School report cards

Core operations (ongoing)

Student Recruitment
Recruit and retain target population of students

Metrics:
17. Enrollment %
18. Free/reduced lunch %
19. Retention %

Talent Development
Attract, develop and retain outstanding teachers and principals and AF staff

Metrics:
20. Acceptance %
21. Retention %
22. Employee performance (PGP results)
23. Employee satisfaction

Finance
Ensure strong financial processes and culture with goal of operating at scale within publicly available funding at NY levels

Metrics:
24. Costs as % of local district funding
25. Clean audits
26. Good budget performance

Learning and Growth (12-18 months)

Network Model
Build seamless support services that scale

Metrics:
27. % of teams meeting goals (First Class Plans)
28. School feedback to central

Organizational Strength
Build next generation management team, organizational structure and processes

Metrics:
29. Number of open positions
30. Forward looking, board approved, organizational charts
31. Succession plan

Facilities
Secure appropriate facilities at affordable cost

Metric:
32. % of schools with secure, affordable, timely facilities

Source: AF Strategic Plan, April 2009, p. 10

Appendix C: AF's Cycle of Highly Effective Teaching

The Cycle of Highly Effective Teaching

ACHIEVEMENT FIRST

Before The Year Starts

Scope & Sequence / Interim Assessment Creation*

Setting Measurable BHAGs for the Year and each IA Cycle

Create Long-Term Plan that Breaks Year Up into Units with Clear Dates

Full Cycle Every 6-8 weeks

Create / Modify Unit Plans

Give Interim Assessment & Create Data-Driven Plan

Goal Setting

Core Values & Constant Learning

Personal Organization & Effectiveness

Student & Family Investment

Lesson Planning

Daily Cycle

Daily Effective Instruction (Tier 1)
Great Aims
Assessment of Student Mastery of Aims
Content-specific Knowledge & Strategy
Modeling/Guided Practice (I/We)
Sustained, Successful Ind. Practice (You)
Classroom Culture
Academic Rigor
Student Engagement
Cumulative Review
Differentiation

Effective Interventions (Tier 2)

* This work is done or facilitated by the curriculum team and edited by teachers and school leaders.

Daily Exit Slip (and quiz every week or two) & Data Analysis

Source: www.achievementfirst.org, Retrieved June 2009

APPENDIX D: The Essentials of Effective Instruction

(1) GREAT AIMS: Rigorous, bite-sized, measurable, standards-based aim(s) are written on the board and reviewed with scholars; the aims clearly drive the activities, not vice-versa.

(2) EXIT TICKET / ASSESSMENT OF STUDENT MASTERY OF THE AIMS:
   a. Exit Ticket / Assessment: There is a systematic way at the end of class to assess every student's mastery of the aim(s) and to diagnose areas of student misunderstanding (most of the time, assessment is through an exit ticket).
   b. Student Mastery: A very high percentage (at least 85% of students) master the aim.

(3) MOST EFFECTIVE & EFFICIENT STRATEGIES to teach the AIM:
   a. Content Knowledge / Right Strategy: The teacher demonstrates strong knowledge of the relevant standards/concepts and uses the most effective and efficient strategy to guide students to mastery; all information conveyed to students is factually accurate.
   b. Pacing & Urgency: The teacher moves students briskly from one part of the agenda to the next; there is a palpable sense of urgency and purpose in the room. Time is held sacred; the teacher spends the appropriate amount of time on each activity and maximizes each minute spent. The teacher sets clear guidelines for how long activities should take and uses timers, time reminders, and countdowns effectively. The class is set up to maximize efficiency, and the teacher is fully planned and prepared to maximize each moment.

(4) MODELING/GUIDED PRACTICE (I/We or We):
   a. Mini-lesson: The lesson includes a clear "think aloud", explicit modeling, heavily guided practice or other form of clear mini-lesson; examples and step-by-step processes are thoughtfully planned and tightly delivered.
   b. Guided Practice / Declining Scaffolding & Guidance: The teacher then leads students through guided practice with declining scaffolding / guidance so that students eventually provide both the answers and the thought process.
   c. Visual Anchor: The mini-lesson is captured (on whiteboard, butcher paper, overhead, and/or scaffolded notes) so that students can reference it during independent practice.
   d. Check for Understanding: The teacher regularly checks for understanding during GP so that students transition to independent practice when they are ready. (A small number of students may need more guided support during independent practice, and this should not hold up the entire class.)
   Note: Although I/We - You is the bedrock of the vast majority of lessons, there may be times when the teacher chooses to start with a short discovery activity, activation of prior knowledge, or some other strategy to lay a conceptual foundation (often in a You - I/We – You format; lessons should end with the We-You and include ample time for successful You time.)

(5) SUSTAINED, SUCCESSFUL INDEPENDENT PRACTICE (You):
   a. Many successful "at bats": Students have ample, successful "at bats" so that they get to practice the aim independently (at least 15-20 min of independent practice). The YOU activity should be at the same difficulty level as the WE activity so that complexity doesn't increase while support decreases. The teacher MOVES around the classroom constantly during independent practice to assess mastery and provide individual help.
   b. Read, Baby, Read: In reading classes, teachers make sure that "nose in text" time is very high and that independent work time has at least a 7:2 ratio of reading to activity/writing/discussing.

Appendix D: The Essentials of Effective Instruction (continued)

(6) CLASSROOM CULTURE
   a. High Expectations, Clear Routines: The teacher sets (with clear *What to Do* statements) and reinforces clear expectations and routines for high standards of behavior consistent with our common picture; with a *Strong Voice*, the teacher sweats the small stuff (e.g. SLANT, no call outs, no laughing at other students' mistakes) and insists students *Do it Again* if not great.
   b. *Joy Factor:* The class is a fun, joyful place where kids are enthusiastic and excited about learning
   c. Positive-Corrective Ratio: The teacher uses *Positive Framing* to correct behavior and narrate class activity; there is a high ratio of positive to corrective comments; the classroom feels like a place where students want to be; students are nice and respectful to each other, and the teacher is nice and respectful to the students.
   d. Students Own It: Students are given the responsibility, tools, and strategies to fix problems they have or created. The teacher resists the temptation to be the sole problem-solver; students who make mistakes must own and fix them.
   e. Teachable Character Moments: The teacher uses key moments in class to explicitly talk about, celebrate, and reinforce character skills; these moments flow naturally from the lesson and are quick and high-impact; the teacher strategically picks examples, texts, and activities that, when appropriate, reinforce the key messages (e.g. going to college, REACH values, etc.).

(7) STUDENT ENGAGEMENT
   a. *100%:* The teacher insists on 100% of students on task with hands consistently in the air (students are either asking or answering questions)
   b. Engagement Strategies: The teacher uses high-engagement strategies (e.g. *Cold Calling*, rapid-fire *Call-and-Response*, mini white-boards, frequent choral responses, non-verbal responses, and/or "everyone writes") to ensure that all students are accountable for engagement; makes it impossible for students to be desk potatoes and simply copy from the board; the teacher limits use of round-robin reading or questioning strategies that engage only one student at a time.

(8) ACADEMIC RIGOR
   a. Teacher Talk–to–Student Work: There is a high ratio of student work to teacher talk with students doing most of the "heavy lifting" of doing the work and explaining their thinking.
   b. Planned, Rigorous Questioning: The teacher plans his/her key questions in advance with a range of questioning – both lower-level (knowledge recall and basic comprehension) and higher-level (application, analysis, synthesis, and evaluation); the teacher regularly uses the *Stretch It* technique – WHY? What does that relate to? How would you apply this?
   c. Top-Quality Oral Responses: The teacher knows that *Right is Right* and refuses to accept low-quality student responses (insists on correct grammar, complete sentences, use of appropriate vocabulary and sufficient detail/rationale (don't settle for so-so); the teacher is a *No Opt Out* champion -- no students are allowed to "opt out" because the teacher cycles back to students who didn't answer.
   d. Top-Quality Student Work: The teacher sets clear expectations and has an accountability mechanism for ensuring all students complete top-quality work; examples of top-quality work are posted for reference and to celebrate great student work.

(9) CUMULATIVE REVIEW: As a part of the lesson and homework routine, students get fast, fun opportunities to systematically and successfully review and practice skills that they have already mastered; standards included in cumulative review are truly review, and the teacher has a clear method of using data to inform which standards to review.

(10)   DIFFERENTIATION: The teacher works to ensure that the needs of every student are met. Especially during independent practice, the teacher can work with some students to provide extra support or enrichment and/or can otherwise vary the volume, rate, or complexity of work that students are asked to complete. (In classes that are grouped homogenously by skill level, pronounced differentiation may be less necessary.)

Source: www.achievementfirst.org, Retrieved June 2009