**DIGITAL ACCESS TO
SCHOLARSHIP AT HARVARD**
DASH.HARVARD.EDU

HARVARD LIBRARY
Office for Scholarly Communication

Learning for Adaptation: Building Teacher Career Pathways at DSST Public Schools

The Harvard community has made this article openly available. **Please share** how this access benefits you. Your story matters

Citation	Fishman, Peter. 2015. Learning for Adaptation: Building Teacher Career Pathways at DSST Public Schools. Doctoral dissertation, Harvard Graduate School of Education.
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:16645016
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Learning for Adaptation:
Building Teacher Career Pathways at DSST Public Schools

Doctor of Education Leadership (Ed.L.D.)
Capstone

Submitted by

Peter Fishman

To the Harvard Graduate School of Education
in partial fulfillment of the graduation requirements for the degree of
Doctor of Education Leadership

May 2015

Acknowledgments

This page may be the hardest to write, because no words can capture my gratitude for the many individuals who have supported me along this journey.

I would like to begin by thanking my Capstone Committee. Monica Higgins, my advisor for all three years at Harvard, you model a level of rigor and excellence, coupled with care and compassion, that inspires and motivates. Thank you for believing in me at every step along the way. Amy Edmondson, thank you for your willingness to make the leap across the river and join my committee; your insights and perspective have deepened this capstone and your research has transformed my beliefs about learning. Christine Nelson, you have been a critical friend, thought partner, and champion. Learning from you has been one of the great pleasures of the residency year. Finally, to my TF, James Noonan, thank you for the close reads, thoughtful advice, and encouragement.

The team at DSST Public Schools is one of the most remarkable groups of people with whom I have ever had the chance to work. Bill Kurtz and Jen Goldstein, thank you for your unending support. To the members of the TCP team – Jacob, Steve, Jake, Coreen, Jeff, Nicole, and Stefan – thanks for learning and collaborating in ways that always kept the focus on classrooms and on students. To the students – the Aspen Challenge team, TCP Dinner speakers, and GVR seniors – you are the reason that I am in education.

I have been blessed to have the support of many mentors and supporters in my career: Claudia Merson, Gil Joseph, Dimitry Anselme, Spencer Blasdale, Jenne Colasacco Grant, Laura Clancy, Jim Peyser. Julie Mikuta and Jordan Meranus have been tremendous thought partners since my first day at NSVF. At Harvard, Adria Goodson, Bob Goodman, Lisa Lahey, Florrie Darwin, Jal Mehta, and Chris Gabrieli have changed the way that I think about leadership and transformation in K-12 education.

Cohort 3, you have taught me more about loving kindness than I can ever express. Prime, thank you for showing me what it means to live and lead with compassion and vulnerability. ADJP, our Google Hangouts, emails, phone calls, and texts propelled me through this year. To my peer coach, Jeremiah Newell, I remain in awe of your wisdom, generosity, and capacity for unconditional love. Thank you. To my Denver partner-in-crime, Landon Mascareñaz, I will be forever grateful for our Sunday afternoons, Wednesday evenings, and Tacos Jalisco. To Brian, JP, and Doannie, let's run Newport every year until our knees give out.

Finally, thank you to my family. Mom, Dad, Annie, and Scott, I am who I am because of you. You taught me to live with simplicity and purpose.

Ariel, you are the source of my daily strength and inspiration. I could not have completed this journey without you. Thank you for keeping me grounded in the present, in joy, in humility, in gratitude.

I suspect that the real moral thinkers end up, wherever they may start, in botany. We know nothing for certain, but we seem to see that the world turns upon growing, grows toward growing, and growing green and clean.

- Annie Dillard

Practitioners know things that policymakers could not, and they use that to adapt policy to their circumstances.

- David Cohen and Susan Moffitt

Table of Contents

Abstract

The past five years have witnessed unprecedented changes in teacher evaluation systems, as states and districts have responded to federal incentives under the 2009 Race to the Top grant competition. Across the country, teacher evaluation has become more frequent, higher stakes, and more explicitly tied to student achievement outcomes. This capstone examines the initial implementation of one such system at DSST Public Schools, a network of nine consistently high-performing charter schools in Denver, Colorado. I describe my role in leading a cross-functional team charged with implementing, learning from, and improving the system in its first full year. I explore the everyday challenges of positioning evaluation as a resource to drive teacher development, arguing that evaluation will achieve lofty developmental goals only if teams of system designers and system users decouple the measurement and development purposes of teacher evaluation and commit to the creation of complementary systems. Leaders can accelerate this process by modeling a developmental mindset and creating conditions for purposeful team learning that leads to system adaptation. This capstone offers important lessons for practitioners and policymakers seeking to position teacher evaluation as a resource for teacher growth and development; for DSST Public Schools in its effort to develop teachers and leaders across its expanding network of schools; and for myself as an educational leader.

Introduction

"The work of building teacher performance management systems is new, promising work and the sector is at the beginning of the learning curve. These systems will be most effective in driving student achievement if leaders commit first and foremost to learning from their efforts and continually refining their design and implementation based on the learning. Approaching this work as dynamic, evolving work will be uncomfortable for some in systems who have often operated in a culture of command, control, linear thinking, and/or compliance. Refinement requires ways for systems to learn from the work in process and to develop flexibility that allow for adaptation."

- Curtis (2011)

Context

Under the Obama Administration's 2009 Race to the Top grant competition, states were encouraged to mandate improvements in local teacher evaluation systems, making them more frequent, higher stakes, and more explicitly linked to student achievement outcomes. According to a 2014 report, the number of states requiring annual evaluations increased from 15 to 28 and the number of states tying evaluations to student achievement increased from 15 to 41 as a result of Race to the Top (Rotherham & Mitchel, 2014). The National Council on Teacher Quality called these changes "a seismic shift rarely seen in education policy" (Doherty & Jacobs, 2013, p. 2).

One of the nation's most ambitious evaluation laws passed during the Race to the Top period was Colorado's Senate Bill 10-191, which requires local education agencies to develop fair and transparent systems of teacher evaluation, at least 50% of which must be based on student outcomes (Senate Bill 10-191). The law directed a newly constituted Council for Educator Effectiveness to produce recommendations for teacher performance standards of "highly effective," "effective," and "ineffective" and revised existing state statutes to make teacher tenure contingent on demonstrated effectiveness (Senate Bill 10-

191, Section 5, 3a). As the bill's author, State Senator Michael Johnston, later wrote in an op-ed with the state teachers union's president, the goal was "to build an effective educator-evaluation system that focused on professional practice and that fostered the academic achievement of every Colorado student" (Johnston & Dallman, 2014). Under SB 10-191, all Colorado public school systems were required to design and begin implementation of new evaluation systems by July 2013.

In designing their evaluation systems, Colorado's charter schools were granted considerable latitude under SB 10-191. Rather than comply with the letter of the new law, charter schools were required only to develop evaluation systems that met the "intent." According to guidance provided by the Colorado Department of Education, meeting the law's "intent" meant (1) evaluating teachers and administrators annually; (2) tying at least fifty percent of the evaluation to student outcomes; and (3) clearly defining "criteria for assigning educators to evaluation categories or otherwise communicating to them about their performance" (Colorado Department of Education, 2014). Importantly, given that teacher tenure laws did not apply to the state's charter schools, personnel decisions were not required to be contingent on demonstrated effectiveness.

Residency Site

This capstone examines the initial implementation of a teacher evaluation system at DSST Public Schools, a network of consistently high-performing charter schools in Denver, Colorado.[1] Chartered by the district, Denver Public Schools (DPS), the Denver

---

[1] The term "high-performing charter school" is used frequently in both the academic and popular literature. Although the field lacks a single, shared definition of the term, it most often refers to a school whose students significantly outperform similar demographics of students in local district-run schools, as

School of Science and Technology (known by its abbreviation, "DSST") was founded as a public charter high school in 2004. In 2010, the school's founding principal, Bill Kurtz, and members of the founding board established DSST Public Schools (hereafter, simply, DSST) as a charter management organization set up to grow a network of middle and high schools across the city. Currently a system of nine schools, DSST plans to grow to twenty-two schools by 2024, at which time the network will be poised to serve approximately 25% of Denver's 6-12 public school population. DSST schools rely upon the district for facilities and some academic and operational supports, but the schools do not follow district curriculum or assessments and teachers and administrators are not part of the local collective bargaining unit. A network Home Office, made up of forty employees, supports over three hundred school-based teachers and staff. Since its founding, DSST has consistently operated some of the highest performing public schools in Denver. In 2013-14, seven of the top nine schools in the city for student academic growth were run by DSST (Carroll, 2014).

Like other school systems throughout the state, DSST responded to passage of SB 10-191 by convening a group of stakeholders to design a new teacher evaluation system. In Summer 2011, a working group of teachers, principals (called "School Directors"), and Home Office staff began to lay the foundation for an evaluation system that would meet the intent of the new law. Up until that point, annual performance reviews of teachers at DSST had occurred on a largely informal basis. School Directors observed and provided oral feedback to teachers at mid-year and end-of-year review meetings. Student achievement data and survey input from colleagues, students, and parents were

measured by achievement on state assessments (see Merseth (2009), Toch (2009), National Association of Charter School Authorizers & Charter School Growth Fund (2014)).

not factored formally into the review, and School Directors had broad discretion to allocate annual raises among their teachers. Teachers were – and still are – at-will employees. Performance reviews informed, but did not dictate, retention decisions for the following academic year.

In 2011, the working group began by defining the characteristics of "Master Teachers," those veteran educators who had the greatest impact on students and school community. The group subsequently mapped a career progression from Novice to Master. Over the course of the next two years, the working group designed a range of measures that would be used to determine a teacher's placement along the career pathway. In compliance with SB 10-191, student outcomes would represent 50% of the weighted measures. The remaining 50% would be made up of a combination of observational data and student, parent, supervisor, and peer surveys. The stated purpose of the evaluation system, called the Teacher Career Pathway (TCP), was "to realize the critical priority" of "identifying, developing, and retaining effective teachers" (DSST Public Schools, 2014).

On paper, TCP differed in a couple important ways from systems designed by many of Colorado's traditional school districts. First, since DSST, as a charter school organization, was not required to tie evaluation categories to definitions of "effectiveness," the designers of TCP built the system around a concept of career progression. Indeed, the "Novice" category was intended not as a commentary on a teacher's absolute effectiveness, but rather as a reflection of an observable stage in a teacher's professional growth. A Novice teacher was not considered "ineffective"; rather, he was simply at an earlier stage in his professional growth than, say, an

9

"Accomplished" teacher was in her development. By contrast, under Denver Public Schools' evaluation system, called LEAP, the lowest category was a "rating" of "Not Meeting," a reflection of absolute effectiveness. The designers of TCP hoped to set an expectation that becoming a Master Teacher was a long-term endeavor, often beginning at Novice and progressing through the stages of the pathway over many years. The evaluation was a measurement about where teachers stood along that pathway.

Second, the designers of TCP chose to tie teacher salaries to the career pathway, believing that significantly higher salaries at the upper ends of the pathway would incent teachers not only to remain in the profession, but also to continue pursuing professional growth opportunities. On a pro-rated basis, the Master Teacher salary was intended to be commensurate with that of a School Director, sending the message that teaching was just as valued as school leadership – a point that was being lost on scores of DSST teachers, who, at the time, voiced an interest in leaving the classroom in order "to advance" in their careers.

By defining evaluation categories and tying compensation to the concept of a career pathway, DSST used its flexibility as a charter school under SB 10-191 to create a system that, the task force believed, would identify, develop, and retain effective teachers. In September 2014, a DSST School Director, who had served as a DPS assistant principal when LEAP was being developed, told me, "It's nearly incomprehensible to me that LEAP and TCP were produced by the same piece of legislation" (School Director conversation, September 16, 2014). LEAP seemed to label teachers based on effectiveness, he explained, whereas TCP aimed to support teachers for

long-term growth. This difference could be traced back to the flexibility that DSST enjoyed as a charter school under SB 10-191.

Despite TCP's stated intent, when I entered DSST as an Ed.L.D. resident in July 2014, there was evidence that teachers were not experiencing the new TCP system as a support for long-term growth. After two years of small-scale pilots, nearly all DSST teachers had been formally observed and received student, parent, peer, and supervisor survey scores in the 2013-14 school year. These inputs generated half of a total evaluation score, which teachers received in May 2014. The other half of the evaluation was pending student achievement data from a variety of sources, including the state and the district. Based on these pending data, career pathway "placements" were to be produced and communicated in Fall 2014. On a May 2014 network-wide survey, teachers had been asked to respond to 85 statements related to their experience working at DSST; these statements ranged from operational to academic.[2] Of the 85 statements, the lowest rated, on average among 170 teachers, was: "In charting my career pathway, the Teacher Career Pathway is helpful." Comments like this were typical: "TCP is, at this point, simply a sophisticated evaluation tool. It really doesn't help teachers guide their career decisions." In the Home Office, there was talk of "TCP losses" – teachers who *left* the school network because they didn't like the way that TCP was being developed.

---

[2] Operational statements included: "I know whom to contact when I have questions about my benefits." Academic statements included: "The time spent norming shared assessments has been valuable in helping teachers on different campuses coordinate how to score interim assessments."

Project

  My strategic project focused on adapting TCP towards its original purpose of serving as a resource for teacher growth and development, even as the new system was still being developed and initially implemented. I entered DSST just as the chief coordinator of the TCP project, the organization's Director of Human Capital, was leaving. For five months, I convened and led a cross-functional development team in the Home Office. This was a group of senior managers and directors with responsibility for processing large amounts of recently-released student achievement data; producing career ladder "placements" based on these data and other inputs; shifting to a new compensation model based on the career ladder; and, most critically, positioning the system as a developmental resource facilitating improvements in instructional practice. The initial "placement," or implementation, coincided with a period of learning and improvement to inform future placement cycles. The new Director of Human Capital co-developed the communication strategy and coached School Directors with me; however, on a daily basis, I had primary responsibility for leading the project, including convening and facilitating all meetings of the cross-functional "TCP team" and coordinating the team's activities (see Appendix A, TCP Team charter, for a breakdown of roles and responsibilities). Given the focus on implementation and learning, I entered the project with the following question:

> *What conditions must be in place for an organization to learn from its work in process and cultivate flexibility that allows for adaptation of a teacher evaluation system?*

Organization

This capstone is divided into three main sections. First, in the Review of Knowledge for Action, I examine research into high-performing charter schools, teacher evaluation, adult development, and team learning, seeking to establish a theoretical and empirical foundation for the where, what, why, and how of my strategic project. The Review of Knowledge for Action concludes with a theory of action for my strategic project. Next, in the Description, Results, and Analysis section, I outline the project itself and evaluate the theory of action, assessing my ability to foster flexibility that would allow DSST to begin to adapt the Teacher Career Pathway towards its intended goal of supporting the continuous growth and development of teachers. By fostering flexibility, I mean building the capacity of the organization to treat TCP as an evolving system, subject to continuous modification and refinement. I argue that, in the absence of complementary systems for support, evaluation may be inoperable and even counterproductive as a tool for teacher learning and growth; a critical task of education leadership, then, is to decouple the measurement and development purposes of teacher evaluation and commit to the creation of complementary, but distinct, systems. Finally, in Implications for Self, Site, and Sector, I examine lessons for practitioners and policymakers seeking to position evaluation as a tool to drive teacher development; for DSST Public Schools in its effort to develop teachers and leaders across its expanding network of schools; and for myself as an educational leader.

Organization

The Review of Knowledge for Action is divided into four sections: where, what, why, and how. First, I look at research into instructional design and human capital trends within high-performing charter schools ("where"), a subset of organizations within the broader K-12 education sector. While elements of my strategic project are generalizable across the sector, parts of this work apply uniquely to high-performing charter schools, particularly given their reliance on a "rare human capital model" (Wilson, 2008, p. 2). Second, I explore recent changes in teacher evaluation ("what"), one of the most widespread systemic reform efforts in K-12 education over the past five years. I seek to understand the theories of change underlying two types of new teacher evaluation systems. Third, I combine the "what" with the "why," using theories of adult development to identify both strengths and potential shortcomings of these new evaluation systems. Based upon this research, I propose that, in order for evaluation systems to become genuinely developmental, educational leaders involved in the implementation of these systems must clearly distinguish between measurement and development purposes. Finally, I review literature on team learning in order to generate insight into "how" I might cultivate flexibility that allows for adaptation of an evaluation system.

1. WHERE: High-performing charter schools

Perhaps no approach to education reform has received greater attention over the past decade than the emergence of charter schools, particularly "high-performing"

schools where predominantly low-income and minority students consistently achieve at high levels, outperforming similar demographics of students at local, district-run schools (Hoxby & Murarka 2009; Toch 2009). The 2014 political firestorm over Mayor Bill de Blasio's decision to reject the co-location of three high-performing charter schools in New York City highlights the salience of this issue in the public discourse about education. While charter schools have demonstrated mixed results nationwide, a group of high-performing schools in several cities across the country have consistently shown outcomes unlike those produced by almost any other academic intervention (Angrist, Pathak, & Walters 2011; CREDO 2013).

Scholars seeking to identify the causes of high-performing charter schools' relative success have pointed to a variety of factors. Fryer and Dobbie (2011) identify five "tenets," which they find explain roughly half of the variation in school effectiveness. These tenets include: (1) extended school day and year; (2) use of data to drive instruction; (3) devotion of high-quality human capital; (4) culture of high expectations; and (5) small group tutoring. Identification of these practices led Fryer and Dobbie to attempt to replicate these conditions in a group of traditional Houston Independent School District schools.

Other researchers have examined the coherence of organizational practices. In her study of five charter schools in Boston, Merseth (2009) argues that high-performing schools are characterized by powerful congruence. "Within these schools, the intense focus on the purpose of their work, the dedication and coordination of people within the organization, and the detailed attention to planning all operate in a coherent fashion at every level" (Merseth 2009, p. 11). Mehta (2013) makes a similar observation regarding

a small group of high-performing charter management organizations: they "recruit mission-aligned teachers, develop knowledge consistent with their pedagogical beliefs, and have clear school norms and cultures consistent with their mission" (p. 287).

Rosenberg (2012, p. 77) borrows a concept from Cohen (2011) to suggest that high-performing charter schools are typified by "an infrastructure of practice." In contrast to the "privatized, idiosyncratic practice" (Raudenbush 2009, p. 172) found across the vast majority of U.S. public schools, infrastructures of practice engage educators in "shared, systematic" work (Rosenberg 2012, p. 109), including shared language for describing instructional practice, common curricula and frameworks, and common expectations for development and improvement. Rosenberg offers an in-depth analysis of the Achievement First network of schools, showing how leaders engage in systemic efforts to codify a vision of instructional quality, develop knowledge, and "foster continuous improvement in individual and collective work" (p. 124). These efforts constitute a nascent infrastructure of practice, a rarity in U.S. schools.

High-performing charter schools are also characterized by what Wilson (2008) calls a "rare human capital" model that demands "nearly heroic efforts by teachers" (pp. 2, 4). Teachers and leaders are often young, well-educated, and high-energy, often working seventy hours or more per week. Hess and Higgins (2009) note, "[T]he most successful charter ventures to date have been boutique-style operations that are extraordinarily reliant on talent and passion, philanthropic funding, and exhausting work schedules" (p. 1). While this model has generated impressive results, as high-performing charter schools expand, they risk staff burnout, insularity, and pressure to "keep doing what they have been doing," according to Hess and Higgins (p. 4). In other words,

success achieved through high coherence and high effort may prove to be unsustainable –

both because staff cannot maintain the effort and because success itself creates few

incentives for individuals and organizations to adapt and evolve.

These patterns may be mutually reinforcing. As more experienced teachers and

leaders burn out and leave an organization, they take with them the capacity to evolve an

"infrastructure of practice" into one that supports continuous, organizational and

individual learning and growth. However, the absence of an infrastructure that not only

supports current practice, but also creates opportunities for organizational and individual

growth may drive experienced teachers and leaders to burn out and depart, thus creating a

vicious cycle where the organizational infrastructure stagnates and individuals feel

unfulfilled. While generally laudatory of Achievement First's infrastructure of practice,

Rosenberg (2012) identifies this vicious cycle as a key concern:

> 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 … 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 (p. 272-3).

The picture that emerges, then, from the research on high-performing charter

schools is simultaneously different and similar to the K-12 sector as a whole. On the one

hand, the presence of coherent infrastructures of practice suggest that educators within

these institutions share a common language and expectations about instructional practice,

anchors against which to ground systems for human capital development. Most K-12

systems lack this basic starting point. On the other hand, teacher and leader turnover

leaves high-performing charter schools in a position that is well-known across the sector

– grasping for a means of systematically fostering the continuous growth and development of all teachers.

2. WHAT: Teacher evaluation

The past five years have witnessed unprecedented changes in teacher evaluation systems (Rotherham & Mitchel, 2014). These changes can be traced back to 2009's *The Widget Effect*, a study conducted by teacher quality organization TNTP, which found that among teacher evaluation systems in four states, less than one percent of teachers received unsatisfactory ratings, with the remainder treated as "interchangeable parts," each equally effective instructionally as the other (Weinberg, et al. 2009, p. 4). The report's authors bemoaned a "national failure to acknowledge and act on differences in teacher effectiveness" and called for new systems that would recognize highly effective teachers and lead to the dismissal of ineffective ones (p. 2).

Rarely has a single study so quickly found its way into public policy. Indeed, a recent appraisal of the report's impact called it a "bombshell" (Di Carlo, 2014). Less than a month after the *Widget Effect* release, Secretary of Education Arne Duncan told members of the National Education Association that the Obama Administration hoped to incite changes in "flawed" teacher evaluation systems that "treat[ed] all teachers like interchangeable widgets" (Duncan, 2009). Weeks later, the Administration released draft guidelines for its Race to the Top grant competition, part of the American Recovery and Reinvestment Act of 2009. Under Race to the Top, states were evaluated on the extent to which they proposed to do five things in the area of teacher evaluation:

> 1. Establish clear approaches to measuring student achievement growth for individual students.

2. Design and implement rigorous, transparent, and fair evaluation systems for teachers.
3. Differentiate effectiveness using multiple rating categories that take student achievement growth into account as a significant factor and are designed with teacher involvement.
4. Conduct annual evaluations that include timely and constructive feedback and provide teachers with data on student achievement growth for their students, classes, and schools.
5. Use evaluations to inform decisions about staff development, compensation, promotion, tenure, certification, and removal of ineffective teachers. (Institute for Education Sciences, 2014).

Although the *Widget Effect* report cited the importance of developing teachers "in the middle of the performance spectrum," the overarching message behind the report was that systems should do a better job identifying high performers and removing low performers (Weinberg, et al. 2009, p. 2). This message rests on the assumption that teacher effectiveness is observable and quantifiable and that focusing attention on the top and bottom ends of a hypothetical teacher effectiveness distribution curve will lead to improvements in overall teaching quality and student learning outcomes. This assumption found perhaps its most dramatic articulation in a 2011 article by Eric Hanushek, whose analyses concluded that the U.S. could close its global achievement gap with higher performing countries like Canada and Finland by replacing the least effective 5 to 7 percent of U.S. teachers with "average teachers" from the middle of the teacher effectiveness distribution curve. While the article was widely critiqued, its basic argument, made by one of the field's most prominent researchers, captured a fundamental assumption of evaluation's proponents – that school systems could "measure and replace" their way toward improved teacher quality (Di Carlo, 2012).

Indeed, *The Widget Effect* reflects what Marzano (2012) calls the "measurement" purpose of teacher evaluation – that systems should be built primarily to distinguish

between effective and ineffective teachers. On their face, evaluation systems built primarily for measurement represent an "exoskeleton" reform in that they fail to foster coherence at the level of classroom practice – it may be *possible* to distinguish between effective and ineffective teaching practice using outcomes that are observable and quantifiable, but that does not mean that system actors *will* make these distinctions or that the process will improve practice (Cohen & Moffitt 2009).

To date, the data are mixed on the effect on student achievement of new evaluation systems that take measurement as their primary purpose. A 2013 study of Washington, D.C.'s IMPACT evaluation system found that "low-performing" teachers who faced dismissal as a result of their evaluation improved their performance in the following year at a higher rate than teachers who did not face dismissal, while financial incentives also improved the performance of the district's highest performing teachers (Dee & Wyckoff, 2013). However, other evaluation systems may be failing to achieve even the minimal goal of distinguishing performance based on measurement. In Pittsburgh Public Schools, 96.9 percent of teachers were rated distinguished or proficient in 2014 under the district's highly lauded evaluation system (Chute, 2014). Across New York State, 94 percent of teachers were rated highly effective or effective in 2012-13 (Harris, 2014). These numbers echo the statistics decried by Weinberg et al. under the old evaluation systems. It seems that perhaps little has changed.

If measurement represents one purpose behind teacher evaluation reform, the second, according to Marzano (2012), is "development." According to Marzano, systems designed for teacher development do not aim to distinguish between "effective" and "ineffective" teachers, rewarding effective teachers and removing ineffective educators,

but rather they help to identify the developmental stage of a teacher and align feedback, coaching, and growth opportunities to that individual's needs. Indeed, as Marzano writes, the goal of an evaluation system designed for development is to "identify areas of strengths and weakness and then systematically begin improving those areas of weakness" (2012, p. 19).

Among high-performing charter schools, the "development" purpose has largely guided the design and creation of new evaluation systems. In 2009, the same year that TNTP released *The Widget Effect*, the Achievement First (AF) network analyzed by Rosenberg (2012) began creation of the AF Teacher Career Pathway, "a systematic, coordinated approach to recognizing and developing great teachers as they progress through five career stages" (Achievement First, 2011). According to Curtis (2011), AF's Teacher Career Pathway rests on an evaluative component, the Teaching Excellence Framework (TEF), which includes four primary elements intended to identify the career stage of a teacher. These elements are:

1. Core values and contributions to team achievement, which are measured through peer and principal surveys.

2. Quality of instruction and planning, measured through observations aligned to AF's infrastructure of instructional practice.

3. Student character development, measured by student and parent surveys.

4. Student achievement, based on a range of internal and external assessments.

The Teaching Excellence Framework produces a placement for each teacher along a five-stage career continuum. Teachers who are new to teaching and to the network enter at

the "Intern" stage. Teachers who are new to the network but not new to teaching generally enter at the "Teacher" stage. Based on two years of strong TEF results, defined by particular cut scores on the framework and a principal's discretion, teachers move to a "strong, stable contributor" stage. With additional years of consistently strong TEF results, teachers move to "Senior" and "Master" teacher stages (Rosenberg, 2012; Curtis, 2011; Achievement First, 2011). Professional development and differentiated responsibilities are intended to align with a teacher's place on the career pathway. For instance, Master teachers receive personal coaching from external instructional "experts," while also serving as coaches to teachers who are earlier on the continuum (Achievement First, 2011). As Curtis notes, "One of the striking aspects of AF's performance management system is the tight alignment of teaching expectations (standards), support, and accountability" (2011, p. 21).

It is difficult to overstate the impact of AF's Teacher Career Pathway on the design of new evaluation systems within high-performing charter schools. In 2010, NewSchools Venture Fund, a key funder of charter schools, highlighted AF's system in a report on "creating talent development systems that drive instructional excellence" (NewSchools, 2010, p. 1). A year later, the Aspen Institute's Education and Society Program held up AF's system as a model for school systems across the nation, saying that the network had successfully codified a vision of effective instruction and aligned evaluation, professional development, and support behind that vision (Curtis, 2011). In 2012, a report from Public Impact and the advocacy organization 50CAN called out the AF Teacher Career Pathway as a model (Doyle & Han, 2012).

Perhaps as a result of this attention, charter school networks sought to emulate the AF system as they responded to teacher evaluation mandates under Race to the Top. For instance, the system at the center of this project, DSST's Teacher Career Pathway, uses a similar set of elements as AF's Teaching Excellence Framework to identify the career stage of a teacher. Based on the framework, teachers are placed on a five-stage continuum, progressing from "Novice" to "Master" teacher. Professional development and differentiated responsibilities also mimic the AF model. The stated purpose of DSST's Teacher Career Pathway is developmental: the system exists "to identify, develop, and retain effective teachers" (DSST Public Schools, 2014, p. 3).

3. WHY: Adult development

A brief review of the literature on systemic approaches to adult development suggests that evaluation systems like the ones created by AF and DSST may need to evolve in order to become genuinely developmental, despite their explicitly avowed purpose of developing teachers.

Before describing systemic approaches to adult development, a definition is in order. By "development," I mean the acquisition of both technical and adaptive competencies. To borrow Heifetz's (1994) distinction between *technical* and *adaptive*, a technical competency might refer to a skill that has been mastered by others. A novice teacher learning how to use Wiggins and McTighe's (1999) "Understanding by Design" framework to create a curriculum unit could be seen as developing a technical competency. Over time, the teacher might deepen his practice by, for instance, using the framework to create multiple differentiated curriculum units for different learners. This

is the view of development most explicitly captured by DSST's Teacher Career Pathway – technical competencies related to instructional practice.

However, "development" also includes the intentional growth of what Kegan and Lahey (2009) call "mental complexity" (p. 12). Mental complexity refers to the ability increasingly to take as *object* that which one previously took as *subject*. At the level of self-authoring to which some, though not all, adults eventually transition, mental complexity means the ability to adhere to an internal belief system, "to self-direct, take stands, set limits" apart from the expectations of one's external environment (Kegan & Lahey, 2009, p. 17). Increasing mental complexity helps individuals confront adaptive challenges, those problems for which no solution readily exists. For instance, a self-authoring teacher trained in the technical skill of backwards curriculum design would respond to a student's learning difficulties not only by examining the alignment between curricular activities and student ability, but also by interrogating his own assumptions and beliefs and seeking to see the world from the perspective of the student. A holistic developmental system would cultivate employees' mental complexity.

Researchers studying systemic approaches to adult development similarly identify two dominant ways of conceptualizing the relationship between people and systems. The first is bureaucratic theory, originally articulated by Max Weber (1978), which emphasizes the importance of control, rules, and procedures. Based on factory models of the late 19th- and early 20th-centuries, bureaucratic theory views organizations largely as hierarchical, where plans and directives issued from above guide the actions and career arcs of employees. A particular view of the human condition informs bureaucratic theory. McGregor (1960) describes this view as "Theory X," the belief that human

beings will generally avoid work and responsibility; as a consequence, supervision and control are necessary for organizational effectiveness. People "develop" at work by being told how to do things differently. Building upon the work of McGregor, Argyris (1971) argues that Theory X leads not only to a top-down, directive style of organizational management, but also to a "soft" approach that emphasizes persuasion and "bought" compliance but is no less pessimistic about human motivation than the "hard" approach of directive management.

Human resource development theory is the second dominant way of imagining the relationship between employees and organizations. This theory views human beings as inherently creative and self-motivated; the role of organizational leaders is to align employee motivation with organizational needs. Surveying the literature on people development, Owens and Valesky (2011) write: "The central mechanism through which the nonbureaucratic organization exercises coordination and control is the socialization of participants about the values and goals of the organization, rather than through written rules and close supervision" (p. 17). McGregor (1960) termed the nonbureaucratic view of the human condition "Theory Y," which sees in all human beings an innate desire to engage in meaningful work.

Those working to develop educators within the American system of public education must contend with the fact that the system reflects the industrial era in which it was born. Indeed, it may be "the starkest example in modern society of an entire institution modeled after the assembly line" (Senge et al., 2000, p. 30). Groups of students are batch-processed according to age, directed where to go and what to learn from the opening bell to the closing bell. A division of labor between administration and

teachers governs the organizational life of traditional school systems. The bureaucratic theory reigns. As Senge (2000) writes, "In a system based on maintaining control, it is the job of the teachers to control the students, the administrators to control the teachers, and the school board to maintain control over the system as a whole" (p. 44).

The measurement view of teacher evaluation takes bureaucratic theory as its starting point. Under the measurement view, teachers are judged on a continuum from ineffective to effective; those in between might be asked to create performance improvement plans to move towards effective. Ineffective teachers are to be counseled out of the profession. Effective teachers are to be retained and rewarded with additional compensation (TNTP, 2010). It is a system based on control.

While seemingly more optimistic about the ability of educators to exercise agency over their professional growth, the developmental view of teacher evaluation nevertheless struggles to move beyond bureaucratic theory. Administrators control most of the levers of the evaluation system – from the observation process to the implementation of surveys and the calculation of student achievement metrics. The developmental view may advocate for teachers to be placed along a trajectory, but it is the administrators, not the teachers, who ultimately determine the placement. In a bureaucratic paradigm based on control and lacking in "shared, systematic" work, this approach may very well reflect a significant improvement on current human capital development practices in school systems. However, it largely fails to account for the cultivation of mental complexity. It rarely builds the capacity of educators to "self-direct." Thus, despite the aspirations of AF and DSST system designers to create resources for teacher growth and development, there is good reason to believe that teachers will experience most evaluation systems as

primarily measurement tools – teachers' self-worth determined by the administrators who maintain control over the system.

A far different picture emerges when one looks at organizations defined by the human resource development perspective. According to Argyris, organizations based on McGregor's "Theory Y" view employees as human beings who bring their whole selves to the workplace, seeking purpose, self-direction, and connection: "Success in the work and the interpersonal contexts are assumed interdependent, with important satisfactions for individuals being achieved within the context of accomplishing important work" (Argyris quoted in Owens and Valesky 2011, p. 19). Similarly, Ouchi (1981) finds a "holistic concern" in successful corporations in Japan and the United States. Ouchi writes, "Relationships between people tend to be informal and to emphasize that whole people deal with one another at work, rather than just managers with workers and clerks with machinists" (p. 79).

A holistic orientation undergirds recent work by Kegan, Lahey, Fleming, and Miller (2014) analyzing the practices of "deliberately developmental organizations" (DDO), companies that are "committed to developing every one of their people by weaving personal growth into daily work" (p. 4). These organizations view performance and individual development as "interdependent" – a developmental culture permits employees to take risks, learn from failure, and collectively improve performance. Kegan et al. find that DDOs engage employees in processes that force self-reflection and public vulnerability. These processes create bonds between team members that lead to the development of technical competencies, mental complexity, and feelings of personal fulfillment.

Kegan et al. (2014) are quick to caution that creation of a deliberately developmental culture is difficult work, involving "a way of life" that requires employees to be vulnerable and to "show themselves at their worst" in a manner uncomfortable to many individuals (p. 7). Indeed, many organizations may choose to create cultures of learning, but stop short of adopting the intensive practices of DDOs given the potential discomfort in store for employees and organizational leaders who pursue the DDO approach. Nonetheless, the lessons from DDOs suggest that when organizations take a holistic orientation and define "growth and development" as not only technical but also adaptive, they increase the likelihood that employees will experience work as developmental, fulfilling, and, ultimately, more productive. Indeed, Kegan et al. argue that organizations focused on both technical skills and mental complexity free employees from the daily demand of hiding their shortcomings and allow them to perform better at their jobs.

A key challenge, then, facing educational leaders involved in the implementation of new evaluation systems is to enlarge their definition of growth and development and exercise flexibility that allows for continuous adaptation of these systems. The remainder of this review will examine *how* team learning might help to cultivate flexibility that allows for system adaptation.

4. HOW: Team learning

Before exploring the process of team learning, it is worth situating my strategic project within a typology of teams. At its core, this project was a process development project aimed at creating a new teacher evaluation system. In seminal work on

development projects in the automobile industry, Wheelwright and Clark (1992) lay out

four "dominant team structures" (p. 10). The exhibit below depicts these four structures.

Adapted from Wheelwright & Clark, 1992, p. 11
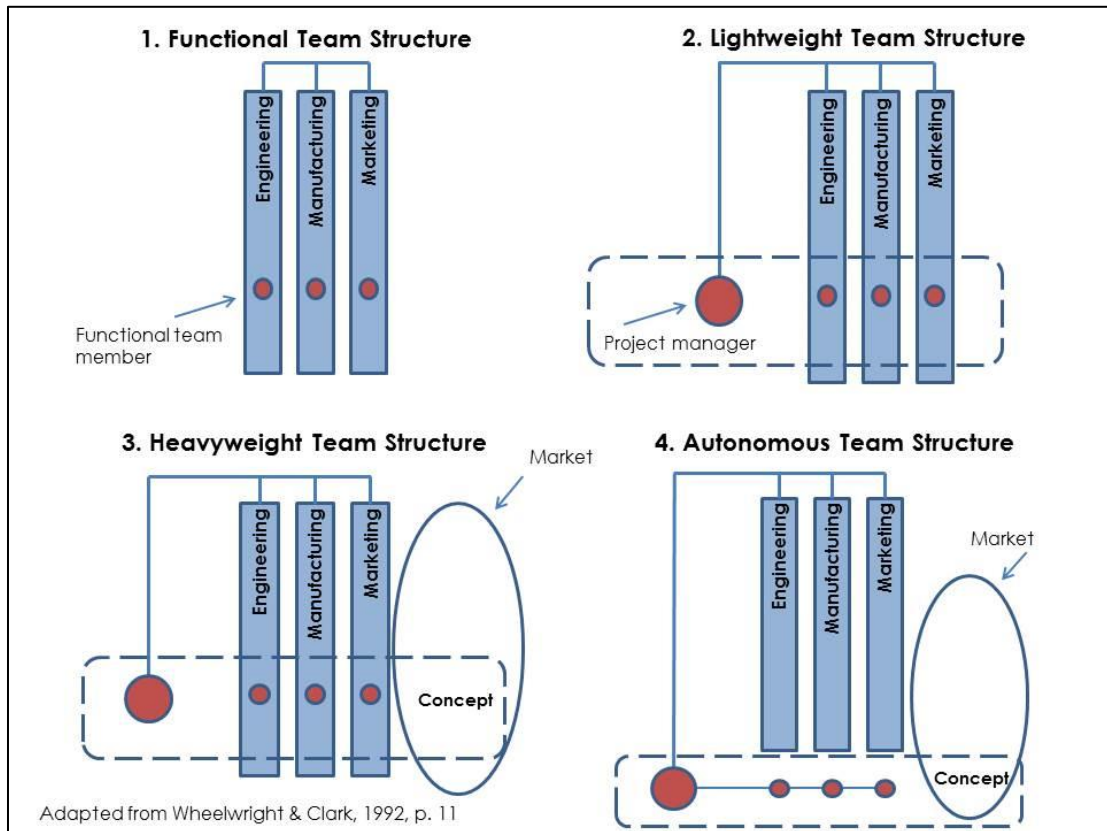
The first, functional team structures, are centered around disciplines, where

individuals work within specialized functions and coordinate development by passing the

project sequentially from one group to the other. At the start of the project, "the entire

development process is decomposed into separable, somewhat independent activities,"

leaving each function to work independently of the others (Wheelwright & Clark, 1992,

p. 10). As such, functional teams work best when the project involves no

interdependencies. Christensen and Kaufman (2008) call functional teams the "default"

mode in most organizations (p. 10).

At the other end of the development process spectrum, autonomous team structures involve a near complete separation of a cross-functional team from the core organization. Representatives from each functional group leave their groups completely for a period of time and assemble to tackle a development project of generally immense interdependency and complexity, where the autonomous team "is not required to follow existing organizational practices and procedures" and where group members report directly to a project lead (Wheelwright & Clark, 1992, p. 14).

If functional and autonomous teams sit on opposite ends of a continuum of cross-functional collaboration, lightweight and heavyweight teams sit in the middle. In a lightweight team, a generally junior-level person serves in a liaison role, overseeing the coordination of functional activities, but exerting little formal or informal influence over the other members of the team (Clark & Fujimoto, 1991). Indeed, "[t]his approach usually figures as an add-on to a traditional functional organization, with the functional liaison person having that role added to his or her other duties" (Wheelwright & Clark, 1992, p. 12). Lightweight teams work best when the project involves "predictable interdependencies" that can be managed with only modest coordination (Christensen & Kaufman, 2008, p. 9).

By contrast, heavyweight teams involve "unpredictable interdependencies," requiring tight integration of diverse functional activities (Christensen & Kaufman, 2008, p. 9). As Christensen and Kaufman (2008, p. 7) write:

> Members bring their functional expertise with them as they join the team. But their mindset must never be to "represent" the interests of their functional group during the team's deliberations. Rather, it is to collectively figure out a better way to knit things together so that the overall project is successful.

Heavyweight teams are led by what Clark and Fujimoto (1991) call heavyweight project managers (HWPMs), individuals who may lack formal authority over members of the cross-functional team, but nevertheless "exercise strong direct and indirect influence across all functions and activities in the project" (p. 255). These leaders serve as both "internal integrators" of the cross-functional team and "concept champions" who "tend to regard coordination and conflict resolution as opportunities to infuse concepts into product designs" (Clark & Fujimoto, 1991, pp. 257, 263). Indeed, in a heavyweight process development project, HWPMs have responsibility for leading the internal team, engaging with users to hone and preserve the process concept, and continually refining the group's work to ensure that the concept is reflected in the actual process (Kuwashima, 2013). As I will discuss in greater detail later in this capstone, I approached my strategic project as an endeavor between lightweight and heavyweight project management – something that I will describe as "middleweight" project management.

In order to be successful at the iterative work of process development, any team, regardless of its structure, must find ways to learn together. Senge (2006) defines team learning as "the process of aligning and developing the capacity of a team to create the results its members truly desire" (p. 218). Contrasted with unaligned team processes where individuals work at cross purposes and energy is "wasted," team learning occurs when "individuals' energies harmonize" behind pursuit of a "shared vision" (p. 217). Based on this definition, one might assume that team learning denotes the absence of conflict or the subservience of individual belief and interest to a higher mission. Neither is true. When teams learn, they engage in productive conflict, surfacing a shared vision

31

that reflects the sum of many personal visions and a way of working that synthesizes the complementary skills and preferred styles of the team's members.

Indeed, according to Senge (2006), team learning has three key dimensions:

1. The need "to think insightfully about complex issues…. Teams must learn how to tap the potential for many minds to be more intelligent than one mind."

2. The need "for innovative, coordinated action…. Each team member remains conscious of other team members and can be counted on to act in ways that complement each other's actions."

3. The need to foster "other learning teams through inculcating the practices and skills of team learning more broadly" (p. 219).

The practices and skills of team learning allow groups to engage in what Argyris and Schön (1978) called "double-loop learning," the ability to interrogate underlying assumptions and make course corrections based on reconsideration of those assumptions. Argyris (1977) famously used the example of a thermostat to illustrate the difference between "single-loop" and "double-loop learning." A thermostat set to turn on whenever the temperature in a room drops below 68 degrees is engaged in single-loop learning – solving a problem based on received information. However, if the thermostat could ask *why* it had been set at 68 degrees and then use judgment to determine whether instead it should turn on at 70, it would be engaged in double-loop learning. Similarly, an HR team that decides to redesign a training based solely on direct employee feedback is involved in single-loop learning. An HR team that asks whether a training is consistent with the

team's shared vision and then redesigns the training accordingly and iteratively is engaged in double-loop learning.

Team learning begins with dialogue, "the capacity of members of a team to suspend assumptions and enter into a genuine 'thinking together'" (Senge, 2006, p. 10). "Dialogue" is distinct from "discussion," a form of conversation where individuals present and defend ideas, probe the ideas of others, and often select a resulting course of action from among the ideas presented. By contrast, dialogue "emphasizes the natural flow of conversation," where assumptions are suspended and the need for action deferred (Schein, 1993). Team learning involves the ability to move seamlessly between dialogue and discussion, and to see dialogue as the natural precursor to the decision-making of discussion. In dialogue, individuals express their personal visions, skills, anxieties and preferred ways of working – with the result being a team that is able to harmonize individuals' energies and operate from a shared vision. Through dialogue, teams learn how to draw upon and complement members' visions, skills, and preferences.

Although dialogue forms the foundation for team learning, many individuals and teams struggle to engage in it. At least three challenges to dialogue and team learning are salient to my project at DSST.

First, most individuals have a natural propensity for avoiding vulnerability, which hinders their willingness to openly share their visions, skills, anxieties, and preferences with groups. As Argyris (1991, p. 103) writes:

> There seems to be a universal human tendency to design one's actions consistently according to four basic values:
>   1. To remain in unilateral control.
>   2. To maximize "winning" and minimize "losing."
>   3. To suppress negative feelings.
>   4. To be as "rational" as possible.

The end result of this "universal human tendency" is that very few individuals willingly practice the vulnerability necessary for effective dialogue and team learning. This is especially true of "well-educated professionals" who have rarely experienced failure and, therefore, "have never learned how to learn from failure" (Argyris, 1991, p. 100). Without experiences of learning from failure, many "bright, capable managers … form a sort of protective shell around [their] deepest assumptions," shutting off the ability to learn in the presence of others (Senge, 2006, p. 233). As noted earlier, many high-performing charter schools, like DSST, employ well-educated, bright, capable, and young managers – the type of individuals identified by Argyris and others as being highly resistant to team learning.

A second challenge to team learning occurs in cross-functional teams, where individuals' "mental models," often reflecting their areas of functional expertise, come into conflict (Senge, 2006, p. 250). Schein (1996) argues that representatives from different functional areas "use different languages and make different assumptions about what is important," leading to confusion and struggle for cross-functional teams seeking to learn together (p. 18). At DSST, for instance, implementation of the Teacher Career Pathway involves the areas of finance, operations, education technology, data analysis, academics and human resources. Bringing these areas together into a single cross-functional team introduces the complicated task of creating a shared language and surfacing divergent assumptions.

A third challenge to team learning involves characteristics of the organization in which the team is embedded. Cameron and Quinn's Competing Values Framework

(1999) identifies four major organizational culture types: hierarchy, market, clan, and adhocracy. Among these four, cultures defined by market competition may be most resistant to team learning because they reinforce the human tendency towards "winning" and often take a resolutely task-oriented (as opposed to relationship-oriented) view of the world, privileging and promoting managers who are "no-nonsense" and "aggressive" (Cameron & Quinn, 1999, p. 23). Indeed, the process of deferring action so that individuals may harmonize energies and cultivate shared vision through dialogue may be particularly anathema to market cultures. And yet, market culture may be the dominant archetype for most high-performing charter schools, which are engaged in a continual competition for students and, to a lesser extent, for philanthropic funding to scale their operations (Chubb & Moe, 1990; Toch, 2009; Rosenberg, 2012). My own decade-long experience as both a participant in and an observer of high-performing charter schools suggests that an intense task orientation, fueled by the competition for students and for funding, contributes both to the creation of coherence at the level of instructional practice, as discussed earlier, and to a resistance to dialogue, whose results are rarely measureable and which may, therefore, be seen as a distraction from the accomplishment of immediate tasks.

In sum, engaging in cross-functional team learning within the context of a high-performing charter school network requires the ability to overcome "defensive reasoning" (Argyris, 1991, p. 100), particularly of well-educated professionals; surface divergent assumptions and create shared understanding; and frame the need for team learning within "the language of the incumbent" (Senge, 2006, p. 299); in this case, that means

explicitly connecting the team learning imperative to the task-oriented culture of the organization.

Edmondson's (2012) research on "teaming" provides insight into what this work might look like in practice. Edmondson describes "teaming" as a fluid process that "includes the interpersonal actions and behaviors required to rapidly collaborate, adjust, and learn" (p. 31-32). In many ways, teaming is akin to team learning: to engage in it effectively, individuals must suspend assumptions, participate in honest, direct dialogue, seek opportunities to complement each other's actions, and embrace failure as an opportunity to learn.

According to Edmondson's "organizing to learn" framework, leaders who create "environments conducive to teaming" practice four leadership actions (2012, p. 83).

1. Frame projects as learning endeavors – as opposed to mere execution activities – by highlighting the interdependence of team members, communicating a compelling purpose, and encouraging collaboration.

2. Create a climate of psychological safety where "people feel free to express relevant thoughts and feelings without fear of being penalized" (2012, p. 77). Leaders must not only *say* that team members are free to express themselves, they must also *act* correspondingly by modeling vulnerability and refraining from penalizing failure.

3. Develop a learning approach to failure by adopting "an inquiry orientation that reflects curiosity, patience, and a tolerance for ambiguity" (2012, p. 184).

4. Facilitate communication across geographic, knowledge, and hierarchical boundaries by framing a shared goal, displaying "curiosity to legitimize sharing

information and asking questions," and providing "process guidelines to help structure collaboration" (2012, p. 212).

As I embarked on my strategic project in July 2014, I believed that its accomplishment would require me to practice these leadership actions: to frame the implementation of the Teacher Career Pathway as a learning endeavor; to create a climate of psychological safety; to adopt an inquiry orientation; and to facilitate communication. By practicing these actions, I hoped to build the capacity of not only the cross-functional TCP team, but also School Directors and teachers to regard the new TCP system as an evolving process that might be improved, modified, and refined towards its intended goal of supporting the growth and development of teachers.

The following theory of action guided my work on the strategic project.

Theory of Action

If I …

- Cultivate a cross-functional development team that is able to learn together in a climate of psychological safety and inquiry;
- Frame the implementation of the Teacher Career Pathway as a learning endeavor and put structures in place to invest key stakeholders (teachers, school administrators, Home Office staff) in improvement; and
- Orient stakeholders around a concept of TCP as a resource for adult development and not merely a tool to measure teacher effectiveness…

Then … I will foster flexibility that allows DSST to begin to adapt the Teacher Career Pathway towards its intended goal of supporting the continuous growth and development of teachers, as evidenced by changes to the TCP system itself and stated perceptions of that system by DSST stakeholders (teachers, school administrators, Home Office staff).

Importantly, I did not enter the project with a preconceived vision about how the TCP system might be improved – a point to which I will return later. Instead, I sought to foster flexibility that would allow key stakeholders to evolve the system. I came to recognize only after completion of the strategic project in January 2015 that I needed to articulate a broad vision that explicitly decoupled the measurement and development functions of TCP and demonstrated how a teacher evaluation system, which measures performance regardless of its design, could nonetheless help teachers grow and develop over time.

II. Description, Results, and Analysis of the Strategic Project

Design of TCP

Entering DSST just as the chief coordinator of the TCP project was leaving, I inherited a system that had been in development for three years. As designed by the 2011-2013 working group, DSST's Teacher Career Pathway rested on four key elements that flowed from the working group's definition about the characteristics of Master Teachers (Appendix B) and resembled the elements in Achievement First's teacher evaluation system:

1. Student achievement. The designers of TCP believed that Master Teachers supported students to achieve excellent learning results, as defined by student performance on a range of both internal and external assessments (e.g. state tests, Advanced Placement exams, internally-developed final exams).

2. Culture contributions. Much of DSST's explicit organizational culture rested upon six core values (respect, responsibility, integrity, curiosity, courage, doing your best). A teacher's ability to uphold these core values was measured through surveys of students, parents, peers, and supervisors.

3. Team contributions. In addition to contributing to culture by upholding core values, DSST teachers were also expected to exercise leadership within their teams (mainly whole school, grade-level, and subject area teams). DSST leaders espoused the notion of an "All Community Teacher," one who was not only a classroom instructor, but also an integral part of a school community. Team contributions were measured through surveys of peers and supervisors.

4. Instructional practice. DSST had defined nineteen "core instructional practices" and created a single rubric, the Teaching Excellence (TE) Rubric, to guide observations of teachers by normed observers. Although observations occurred throughout the year, only two high-stakes observations were scored and factored into a teacher's evaluation. Data from a student course survey also factored into the evaluation.

According to the TCP design, survey and observation data were weighted to produce a single "Culture, Team Contribution, and Instructional Practice" (CTI) score, which made up 50% of a teacher's total evaluation. In compliance with Senate Bill 10-

191, student achievement data accounted for the other 50% (Appendix C). When I joined DSST in July 2014, the organization was awaiting achievement results from the Colorado Department of Education, Denver Public Schools, and the College Board, the non-profit organization that administers Advanced Placement (AP) exams. Once all student achievement results were available, they were to be analyzed and each teacher's results were to be assigned one of three student achievement "levels," which were defined relative to historical performance of DSST students: Level A was below expected; Level B was expected; Level C was above expected. These levels were to be cross-walked on a matrix with a teacher's CTI score to produce a career pathway placement, which corresponded to a new salary and, theoretically at least, to new responsibilities within schools (Appendix D).

The stated purpose of TCP was to "identify, develop, and retain effective teachers." However, when I joined DSST, initial evidence suggested that the system was not living up to its purpose. After many years of enjoying staff retention rates over 80%, retention had plummeted to below 67% following the 2013-14 school year. On end-of-year surveys and in exit interviews, many departing staff members cited TCP as a reason for their departure. One teacher wrote: "TCP is causing an incredible amount of stress on our campus. We are under intense scrutiny for every aspect of our jobs, and the pressure to perform both in the classroom and on tests is so high that teachers are extremely unhappy. It doesn't feel like a pathway."

1. Description of Strategic Project

My strategic project focused on adapting the Teacher Career Pathway in the midst

of its development and initial implementation. The project had three main strands,

aligned to my theory of action. I summarize the three strands below and note key

activities by phase of the project. The following section describes each strand in greater

detail.

		Key activities		
	Theory of Action	**July – August**	**Sept. – Oct.**	**Nov. – Dec.**
1	**If I** … cultivate a cross-functional development team that is able to learn together in a climate of psychological safety and inquiry;	• Convene cross-functional development team • Set vision • Organize to learn	• Weekly meeting to problem solve collaboratively • Ongoing dialogue • Production of pathway "placements"	• Team debrief and reflection process • Planning for continuous improvement based on feedback data
2	Frame the implementation of the Teacher Career Pathway as a learning endeavor and put structures in place to invest stakeholders in improvement;	• Presentation to new and returning teachers at August PD days	• Invest School Directors in learning process by including them in collaborative problem solving	• Organization-wide message and survey sent to all employees • Feedback sessions at schools • One-on-one meetings
3	Orient all stakeholders (teachers, School Directors, Home Office staff) around a concept of TCP as a resource for adult development and not merely a tool to measure teacher effectiveness…	*None*	• Coaching of School Directors to use "placement" conversations as development opportunities	• Dialogue within network-wide Leadership Team and Home Office's Schools and Academic Leadership Team (SALT)

Strand #1: Cultivation of a cross-functional development team

Guided by my theory of action, I entered DSST with the belief that team learning could serve as a mechanism for system adaptation, driving the evolution of TCP. However, before I could convene a team, I needed to diagnose the current state of TCP teamwork. Between 2011 and June 2014, the outgoing Director of Human Capital had led the project, running it largely through the organization's existing functional teams. Although a group in the Home Office met semi-regularly to discuss the project and set an overall vision, I discovered that members of the team did not generally exercise shared ownership to "collectively figure out a better way to knit things together so that the overall project [was] successful" (Christensen & Kaufman, 2008, p. 7). Instead, the Director of Human Capital communicated tasks to functional directors, who distributed assignments among their teams. According to one member of the group, "[The Director of Human Capital] did everything. I wasn't thinking about TCP beyond my deliverable. My deliverable came from her" (S. Coit, October 28, 2014).

The sense of the outgoing Director of Human Capital "owning" the project and delegating tasks through functional teams was visible. During my first official week with the organization, I attended a network-wide Leadership Team meeting where the outgoing Director, in her final week, single-handedly led the assembled Home Office Directors and School Directors through a ninety-minute training on the "key messages" that directors ought to deliver to their teams about TCP. Her leadership seemed to have provided steady vision and consistency throughout the design stages of TCP; with her transition, other directors expressed unease about the absence of a single owner for the new evaluation system.

In late July, I set out to cultivate a cross-functional team that would exercise shared ownership over not only key development tasks, but also the overall concept of TCP as a career pathway. Enlisting the support of the new Director of Human Capital, I invited both the four Home Office staff members who had met semi-regularly from 2011-2013 as well as the organization's Director of Finance, Director of Schools, and Senior Manager of Professional Development to join a new, cross-functional "TCP team" that ultimately included eight Home Office directors and senior managers in the areas of Academics, Operations, Finance, Data, Technology, Professional Development, Instruction, and Human Capital (Appendix A). My rationale for convening this team was two-fold. First, I believed that the project involved unpredictable interdependencies that no single team member, especially myself as a new member of the organization, could anticipate. Second, I thought that if the team owned the concept of TCP as a developmental career pathway, we would have a higher likelihood of infusing it across all aspects of the project – from the way student achievement data was presented to teachers to the content of the instructional rubric.

At our first meeting, I guided the team through an exercise to create a vision for the way we would work together. Team members agreed to the following:

> As a TCP team, we strive to:
> - Continue to build, modify, and improve TCP from FY14 to FY15 so that the entire DSST team has a resource and tool to support ongoing teacher development across the network.
> - Communicate clearly and consistently about design and key decisions, speaking with one voice about all things TCP.
> - Work collaboratively to understand individual responsibilities as well as their fit within the larger system.
> - Continue to leverage internal team member strengths and foundational work of the development and pilot phases in order to accomplish the above.

This vision statement recognized the work that had been accomplished by individuals on the team in the years leading up to July 2014, while also suggesting an important shift in the group's efforts. The third bullet reflected this shift most succinctly: rather than operating as a collection of functional representatives, we would seek to "work collaboratively" and to understand interdependencies.

From late July to November, I convened and facilitated meetings of the TCP team at least weekly both to problem solve about development tasks and to discuss opportunities for continued improvement of TCP. Early on, we generated a series of philosophical questions that had not yet been answered about the design of the system, opening space for team dialogues, which took place throughout the project. Indeed, in structuring each agenda and leading each meeting, I tried to create opportunities for diverse viewpoints to be expressed and for listening to occur, while simultaneously maintaining the team's focus on the tactical work of implementation. My hope was for this team not only to produce fair and accurate placements – simply ensuring that the right student achievement data were linked to the right teachers and that teachers received an accurate change in compensation were not simple tasks – but also to commit to TCP's continual improvement, not merely taking as given the system as it was originally designed by the TCP working group.

An example may help illustrate how the team operated. One concern about the use of achievement in a teacher evaluation system is that a single teacher's impact on student outcomes may vary from year to year. In one frequently cited example, 31% of New York City English teachers who had student achievement scores in the bottom quintile of teachers in 2007-08 had scores in one of the top two quintiles a year later

(Corcoran, 2010; Green, 2014). TCP mitigated for this concern by incorporating up to seven different student achievement measures into a single teacher's student achievement "level," thus reducing the impact of variability in any one measure. Still, if DSST had access to multiple years of student achievement data, shouldn't that data be factored into an understanding about a teacher's impact on student learning? I framed and facilitated discussion of this question over several meetings in August 2014. Although the manager of data and assessment had functional expertise in this area, I told the team that all voices mattered. Indeed, it might be the Director of Operations who had the key insight that moved the group forward.

As a result of this dialogue, the TCP team elected to create an "automatic review process" whereby a committee of School Directors and Home Office academic leaders reviewed multiple years of student achievement data before a final data "level" was determined. Consideration of an automatic review process pushed the team's thinking. During one conversation, the Senior Manager of Education Technology said, "I've been looking at this as how to defend the system and not admit imperfection. What you're suggesting – that we admit imperfection and therefore introduce a human element – is a different way of looking at things I hadn't considered" (J. Firman, August 28, 2014).

Strand #2: Communication about TCP as a learning endeavor

The second strand of my project involved communication about the implementation of TCP as a learning endeavor. If we were going to avoid succumbing to bureaucratic theory – where evaluation is done by administrators to teachers – we had to find a way to frame TCP's implementation as learning and to invest others in the learning

45

process. Fortunately, I was able to build on a strong foundation. During the design and pilot phases of TCP, teachers had worked together to craft survey content, refine rubric design, and provide input on student achievement measures through focus groups and working groups. Subsequently, most of the key components of TCP had been approved by DSST's School Directors and staff had engaged in extensive norming on survey and observation components.

Building on this foundation, I tried to send a clear message that improvement was not confined to the design and pilot phases and that, as an entire DSST team, we were engaged in a learning process that would extend beyond TCP's initial implementation. Communication occurred in three phases.

First, in August 2014, I helped to design presentations about the components of TCP for both new and returning teachers. School Directors delivered these sessions to teachers during professional development days that preceded the school year. As part of this process, the Director of Human Capital and I chose to include several messages about the ongoing evolution of TCP. For returning teachers, we included detailed explanations about the feedback that had been received from teachers during Spring 2014 focus groups and the specific changes that had been made to the design of TCP as a result of this feedback. Most of these changes involved incremental improvements to the observation process and the wording of survey questions. Post-session surveys indicated that returning teachers valued seeing feedback put immediately into action. One teacher wrote that she appreciated "the transparency in the explanation of changes, knowing that teacher voices have been heard!" Another teacher said: "I loved that admin was upfront and transparent about the changes to TCP."

In addition to describing these changes, the presentation also included an explicit invitation to both new and returning teachers to continue to share feedback about the system. We gave teachers a single email address to send their thoughts and crafted a four-sentence summary of the system that concluded with the line: "TCP evolves and improves over time based on feedback from teachers and leaders." Thereafter, we included this line in every overview of TCP. When asked what he had learned from the August session, one teacher wrote: "This is an evolving process."

The second phase of communication occurred between September and October 2014, during the height of initial implementation after we had received student achievement data from the district and state. During this phase, we focused less on engaging DSST's teachers and more on investing School Directors in the learning process. Specifically, as the TCP team began to analyze student achievement data from the prior school year and produce initial placements, we were able to discuss improvements based on actual examples involving teachers.

For instance, the initial data showed that teachers with prior teaching experience would place at Novice, raising important questions: what did it mean to be a Novice teacher? What was the expected growth trajectory of a Novice? The Director of Human Capital and I shared responsibility for communicating about TCP with School Directors and between September and October one of us met individually with each of the nine School Directors to discuss the case of Novice teachers. Through these conversations and in the midst of initial implementation, School Directors elected to change Novice to a one-year placement, meaning that Novice teachers would be re-evaluated the following year (in contrast to other stages on the pathway, where a placement "held" for two years).

This change reflected a belief that a teacher should move more quickly through the

Novice stage than he or she would move through other stages on the pathway. By

engaging School Directors in a process of learning from the initial data and making

adaptations in real time, I sought to cultivate and model flexibility, a willingness to

modify and refine the system as we gained new information.

The final phase of communication involved formal opportunities for teachers and

leaders to offer feedback on the system. In mid-October, I drafted a message to the entire

organization from our CEO, Bill Kurtz. Released on the day that we completed the first

cycle of TCP, the message read, in part:

> TCP is still evolving. I have appreciated the many positive and
> constructive pieces of feedback we have received from you over the last
> year. We are a learning organization that embraces the opportunity to get
> better each day. We have work to do to ensure that TCP is a great
> resource to help drive teacher growth through coaching and feedback,
> differentiated professional development and responsibilities, and
> placement. This is just the beginning. Over the next two weeks, teachers
> will be receiving a TCP feedback survey that we strongly encourage you
> to complete, as well as information on feedback sessions for both TCP and
> professional development.

Releasing this message on the same day that we completed placements reflected a

commitment to learn from the first cycle of implementation and a desire to continue to

evolve the system. A feedback survey followed ten days later. In November, I organized

a series of four feedback sessions across DSST's campuses. Scheduled for one hour,

three of the four sessions each lasted nearly two hours. In addition, sharing responsibility

for communicating about TCP with teachers and leaders, the Director of Human Capital

or I met one-on-one with any teacher or leader who wanted to provide feedback outside

of the formal sessions. At least a dozen teachers and leaders took us up on the offer.

Through each phase, I sought to enroll the "users" of the system – school leaders and teachers – in the task of reflecting upon and suggesting changes to the system itself, moving it closer to its original purpose of supporting teacher growth and development.

Strand #3: Orientation of stakeholders around TCP as a resource for adult development

In order for TCP to become genuinely developmental, stakeholders needed concrete practices that would connect the system to teacher growth and development. I was slow to recognize this fact in Summer 2014. Throughout July and August, I spoke about the developmental purpose of TCP without knowing myself what it would look like for this system to be developmental not only on paper, but also in practice. Part of my challenge was that I was not in schools regularly – outside of a couple weeks of summer school, schools were not in operation during the summer months – leaving me with little context for how development could occur among teachers and leaders. I also assumed – incorrectly, it turned out – that stakeholders had a vision; I simply needed to foster the organizational dexterity that would allow this vision to come to life.

As September approached, I became aware that the organization lacked a shared vision for TCP as a resource for teacher growth and development beyond the basic career pathway outlined by the working group. Specifically, *how* was the system going to foster a teacher's movement from Novice to Developing, Lead to Master, and so on? I may have been cultivating dexterity among the TCP team and framing implementation as learning, but most stakeholders viewed this work as improving the ability of the system to *measure, not develop*. Looking back on the design and pilot phases of TCP, the Senior Manager of Professional Development, who had been a School Director until June 2014,

told me: "I remember so much more time about [the evaluation's percentage breakdowns] and salary ranges versus this is going to be the system that drives you to become the best teacher you can be for our kids.... That made it feel like evaluation and compensation. Maybe there was some conversation around development and growing. It was way less than it needed to be" (J. Osborne, October 30).

In September, I saw an opportunity for School Directors to begin to construct a vision of the system as a resource for development. In order to communicate final pathway placements to teachers, School Directors were to have one-on-one conversations with each of their teachers. Some School Directors viewed these conversations as simply procedural: teachers wanted to know the pathway placement, the School Director would communicate, and the whole conversation would be over in five minutes. In other words, the placement conversations were to be structured as purely evaluative. I reframed these placement conversations as pivotal development opportunities, seeking to build the capacity of both teachers and leaders to understand and own the system as a tool for growth. Critically, these conversations offered a chance for teachers to grapple with the TCP data, to reflect on their strengths and areas for growth, and to consider the career pathway within the context of their professional aspirations.

In the weeks leading up to October 2014 placement conversations, I met with each School Director at least once and sometimes three or four times, coaching him or her on strategies for the conversations. I also created agenda templates, video examples, and other tools to support School Directors (Appendix E). At first, I resisted the idea of creating tools for School Directors, seeing the idea as overly prescriptive and directive. However, when pressed by the Chief of Schools and Director of Schools, the two

members of the Home Office with ongoing responsibility for managing and coaching School Directors, I agreed that by providing actionable tools that could be easily modified and the coaching to support this modification, we could create space for each School Director to construct his or her own vision and practices that felt authentic to his or her leadership style and team. My one overarching directive was that School Directors should take an inquiry stance, asking questions of teachers in order to promote reflection and articulation of improvement steps. Eighty percent of the time in placement conversations was to be devoted to inquiry; twenty percent was to be focused on communication about the actual evaluation measures and changes in compensation.

Whereas embedding a developmental opportunity into placement conversations was largely opportunistic, I tried to create deliberate space to construct shared vision in the weeks following completion of the placement cycle. In November 2014, I led a multi-part debrief process with the TCP team that focused on dialogue about the developmental intent of TCP. This dialogue extended into meetings of the School and Academic Leadership Team (SALT), a weekly meeting that involved not only some members of the TCP team, but also DSST's Chief of Staff, Chief of Schools, and CEO. By late November, stakeholders in the Home Office agreed that more time needed to be devoted to building a shared vision for TCP as a resource for adult development.

However, creating space to build shared vision ran into the very immediate challenges of operating schools. In November, a twenty-four year-old teacher passed away unexpectedly at one of the middle schools, causing incredible pain and grief among staff and students. A School Director returned from maternity leave at a second school, replacing an acting School Director and producing a delicate leadership transition.

Student and staff culture struggled at a third school, where each member of the school leadership team was in his or her first year on the job. Grand jury decisions in Ferguson, Missouri, and Staten Island, New York, led to student walk-outs at all of DSST's high schools. These challenges required time and attention from stakeholders in the Home Office and from leaders in schools. For the monthly network-wide Leadership Team meeting in early December, I planned a series of conversations intended to build shared vision around the developmental purpose of TCP. However, in the days leading up to the meeting, these conversations were scaled back to create time for discussion of immediate challenges; on the day of the meeting, two School Directors were absent, as they managed student walk-outs on their campuses.

By late December, School Directors and Home Office staff voiced support for what I termed a "pivot" from measurement to development, but we had not yet defined what this pivot would look like.

2. Results

Using the three strands of the strategic project, this section summarizes my results. Additionally, I look at the intended outcome of the theory of action: adaptations of the Teacher Career Pathway, as evidenced by changes to the system itself and stated perceptions of the system by DSST stakeholders. I use four sources of data to assess the results of the project:

1. Hour-long semi-structured interviews with seven members of the cross-functional TCP team. The eighth member of the team was on maternity leave from August through December.

2. Data from four organization-wide surveys:

   a. End-of-Year Survey (May 2014) that asked the four questions about TCP. This survey was sent to every employee in the organization. Questions about TCP were asked only of teachers. 170 teachers responded.

   b. TCP Presentation Feedback (August 2014). This survey was given at only five schools and surveyed only teachers. In total, 104 teachers responded.

   c. TCP Survey (October 2014) that I designed and included roughly twenty questions. This survey was sent to every employee in the organization. In total, 157 employees responded: 123 teachers, 19 school leaders, and 15 Home Office staff. Some TCP questions were asked of every employee, while other questions were asked only of teachers placed on the pathway in Fall 2014.

   d. Mid-Year Survey (January 2015) that asked six questions about TCP. This survey was sent to every employee. Questions about TCP were asked of both teachers and school administrators. 219 teachers and 35 school administrators responded.

In order to orient the reader, the timeline on the following page provides an overview of surveys and key placement events between May 2014 and January 2015.

Timeline of Organization-Wide Surveys
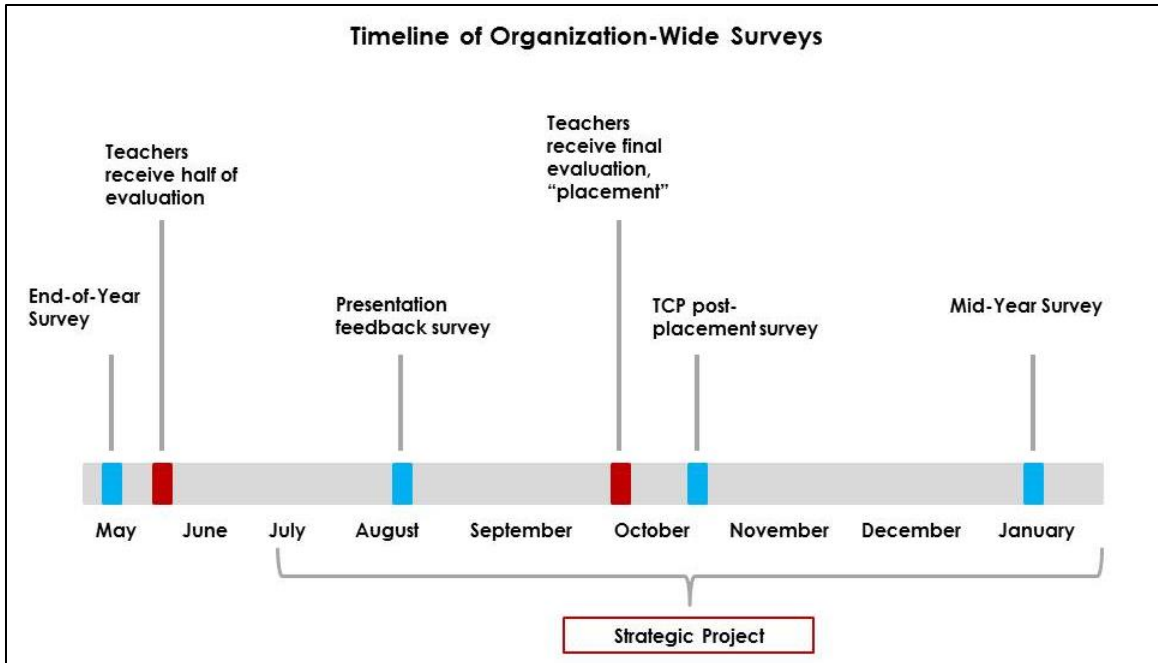
3. Meeting notes and agendas from weekly TCP team meetings, monthly Leadership Team meetings, and dozens of one-on-one and small-group conversations.

4. Observation of events and actions taken by individuals and teams between July 2014 and January 2015.

The chart on the following page provides a snapshot of my results. As I will explain, results from the first strand of the project were mostly positive. Results from the second and third strands were more mixed. Taken together, however, these efforts produced some early indications of a system that was capable of flexibility and adaptation.

	Theory of action	Success to date	Key results
1	**If I** … cultivate a cross-functional development team that is able to learn together in a climate of psychological safety and inquiry;		• Team members reported evidence of psychological safety and learning • Successfully completed 3 of 4 elements of Teacher Career Pathway implementation
2	Frame the implementation of the Teacher Career Pathway as a learning endeavor and put structures in place to invest stakeholders in improvement;		• Belief in implementation as learning was high among Home Office staff and school administrators, mixed among teachers • Investment in improvement process was high among Home Office staff and school administrators, low among teachers
3	Orient all stakeholders (teachers, School Directors, Home Office staff) around a concept of TCP as a resource for adult development and not merely a tool to measure teacher effectiveness…		• Home Office staff and school administrators were beginning to speak about TCP in terms that positioned it as a developmental resource • Teachers mainly oriented towards TCP as a high-stakes measurement tool
	Then…	**Success to date**	**Key results**
	I will foster flexibility that allows DSST to begin to adapt the Teacher Career Pathway towards its intended goal of supporting the continuous growth and development of teachers, as evidenced by changes to the TCP system itself and stated perceptions of the system by DSST stakeholders.		• Four improvement processes underway in January 2015, each intended to effect substantive changes to the TCP system itself • Perception of TCP improved among less experienced teachers who received large raises on account of the system. Perception did not improve among more experienced teachers and among less experienced teachers who did not receive large raises

<u>Strand #1: Cultivation of a cross-functional development team</u>

There are two ways of viewing results from the first strand of my strategic project. First, I look at team process: did we learn together in a climate of psychological safety and inquiry? Second, I look at performance: what did we accomplish as a cross-functional team?

Team process: effectively established a climate of psychological safety and inquiry. To understand whether psychological safety and an inquiry orientation existed on the team, I asked each member of the TCP team to respond to three open-ended questions during interviews in late October:

- To what extent did we act in alignment with our team goals?
- What was a time when the team worked really well together?
- What was a time when we struggled to get something done as a team?

Next, I coded responses for evidence of psychological safety and inquiry by looking at whether team members (i) named an error without assigning blame to individuals; (ii) cited group dialogue about errors; and (iii) described collaborative activity to address errors. I felt that these were important indicators of psychological safety because they captured the extent to which "frank conversations and public missteps" occurred on the team (Edmondson, 2012, p. 118). A summary is below.

	Evidence from individual, semi-structured interviews		
Team member	Named an error without assigning blame	Cited group dialogue about errors	Described collaborative activity
Director of Human Capital	X	X	X
Senior Manager of Data & Assessment	X	X	X
Senior Manager of Education Technology	X	X	X
Director of Operations	X		X
Director of Finance	X		X
Director of Schools	X		X
Senior Manager of Professional Development	X	X	X

As the chart indicates, every member of the team was able to name a human error that had been brought to the group; none of these team members assigned blame for that error to an individual. These errors ranged from an oversight in the data analysis to fraught coordination with an external vendor. Four of seven team members cited group dialogue, which I framed and facilitated throughout the fall, as key to understanding and gaining multiple perspectives on each error. Among the three team members who did not cite group dialogue, other factors, like data analysis, were used to explain the primary way that problems were solved. Whether team members cited group dialogue or data analysis as the primary means of solving problems, all seven team members described the group's activity as highly collaborative. "There came to be a protocol for solving problems," said the Senior Manager of Data and Assessment (S. Coit, October 28). If left unresolved, many of these problems could have undermined the integrity of the system, leading to delays and perceptions of unfairness. However, team members proved willing to admit

errors and work together to solve problems. "There were a few opportunities where we could have gone into panic, reactive mode," noted the Director of Human Capital (J. Goldstein, October 28). Instead, team members came to trust each other to find measured solutions.

Trust seemed to derive from a sense of shared ownership. Whereas the former Director of Human Capital controlled nearly all aspects of the design and pilot phases, I imbued the team with collective responsibility for TCP. "Before I wasn't part of the decisions that got made," said the Senior Manager of Data and Assessment, "now I feel shared ownership of TCP. I will be able to and will be inclined to say my opinion" (S. Coit, October 28). "One of the things that I loved about the team," explained the Director of Operations, "is to see complementary skillsets leveraged to a collaborative team, big picture project that touches every staff member" (J. Roddy, October 30). The sense of shared ownership was echoed across the TCP team. The Senior Manager of Education Technology said: "This is one of the shining examples of true, cross-functional collaboration" (J. Firman, October 29).

Team performance: successfully completed 3 of 4 elements in Teacher Career Pathway implementation. In order for team learning among the TCP team to serve as a model and motivation for flexibility and learning for others in a highly achievement-oriented organization, the team had to demonstrate results. As described earlier, the TCP team's responsibilities included (i) processing large amounts of recently-released student achievement data; (ii) producing career ladder "placements" based on these data and other inputs; (iii) shifting to a new compensation model based on the

career ladder; and (iv) positioning the system as a developmental resource facilitating improvements in instructional practice.

By the end of October 2014, the team had achieved the first three outcomes. All 99 teachers who were slated to be "placed" on the career pathway had received (i) their placements, (ii) a detailed explanation of student achievement data, and (iii) new compensation, reflected in a supplemental payroll on October 17. Acknowledging the shortcomings of using student achievement data to identify a teacher's skill, we had encouraged teachers to request a review of any data that they considered unfair or biased. Only four teachers submitted requests, suggesting that teachers understood and felt respected by the process. On the October 2014 survey, one teacher wrote: "Where I was on TCP and why I was at any given step was clear to me and, on the whole, fair." The Director of Schools, who had served as a School Director until June 2014, attributed this success to the TCP team, telling me: "The collaboration of the team that you put together – having everyone involved – was extremely impactful. There were a lot of misses that could have happened without the team being there" (S. McVoy, October 30).

In processing the data and setting cut scores, the team also avoided the tendency to inflate scores and to place every teacher at the top of the pathway. As noted in the RKA, many new teacher evaluation systems have rated upwards of 95% of teachers at the top of effectiveness scales. Under DSST's career pathway, only nine teachers (9%) placed at the Master stage and only eighteen teachers (18%) placed at the Lead stage (Appendix F). Throughout August and September, the TCP team and, later, the SALT team and network-wide Leadership Team engaged in hours of dialogue about cut scores and the "automatic" reviews described earlier. These conversations were pivotal in

59

producing placements that most stakeholders felt reflected the career stages of DSST's teachers.

Progress on the fourth outcome – positioning the system as a developmental resource – was mixed, as will be discussed in detail later. However, within the team, we came to share a common belief that TCP was a resource for both measurement and development, but that the development purpose should be dominant. The Director of Finance said, "In our smaller group, we were really cohesive about what TCP was" (C. Miller, October 31). In convening the cross-functional team, one of my goals had to been to infuse the idea of development across all aspects of the project. By November, all of the functional representatives on the team were talking about development, not measurement. As the Director of Human Capital said, "We were misaligned at the start, but we've moved towards a healthy place" (J. Goldstein, October 28).

Strand #2: Communication about TCP as a learning endeavor

The second strand of the project focused on framing the endeavor as learning and investing stakeholders in improvement. Again, there are two ways to consider results. First, did stakeholders view implementation as a learning endeavor? Second, were they invested in the process? I found that both belief and investment in TCP's implementation as a learning endeavor varied considerably by stakeholder group.

Belief in implementation as a learning endeavor was high among Home Office staff and school administrators, but mixed among teachers. On network-wide surveys in October 2014 and January 2015, returning staff members were asked to rate their level of agreement with the statement: "I believe that TCP will continue to improve

over time as a resource to support teacher growth and development." This question was intended as a reflection of belief in TCP as a continuously evolving system. Keeping with DSST's historical survey methodology, responses were on a point scale from 0-70, with 0 being "not at all true" and 70 being "definitely true." The results appear below. On both surveys, school administrators (School Directors, Deans, Managers) rated the question nearly twelve points higher on average than did teachers. These results are consistent across both survey administrations. Similarly, on the October 2014 survey, Home Office staff members rated the question on par with ratings of school administrators. (Due to an error in the survey's display logic, the question was not asked of Home Office staff members on the January 2015 mid-year network survey). These results suggest that Home Office staff and school administrator belief in TCP's implementation as a learning endeavor was higher than that of the network's teachers.

"I believe that the Teacher Career Pathway will continue to improve over time as a resource to support teacher growth and development."				
	Oct. 2014 Survey		Jan. 2015 Survey	
Stakeholder group	Average	N	Average	N
Returning Home Office staff	64.4	10	--	--
Returning school administrators	62.86	14	61.96	26
School Directors / Associate School Directors	67.75	4	66.64	11
Returning teachers	50.11	62	50.59	108

However, overall averages do not tell the full story. In order to produce a more nuanced picture of teacher responses to implementation-as-learning, I cut the data by a variety of indicators, including (i) number of years of teaching experience; (ii) number of years with DSST; (iii) subject area; (iv) school; (vii) "placement" on the pathway; and (viii) size of compensation increase resulting from placement in Fall 2014. I found that one group of teachers expressed a much greater belief in TCP's continuing evolution than

did others. As a result of TCP's implementation in Fall 2014, many teachers received

standard raises between 0% and 3.66%; some teachers, however, received extraordinary

raises between 5% and as much as 40%. According to mid-year survey data, teachers

who were earlier in their careers and received these large raises were much more likely

than their peers to express belief that TCP would continue to improve. This group of 29

teachers rated the question 60.38 out of 70 compared to between 46.09 and 47.81 for all

other groups. The chart below demonstrates the difference. Importantly, more veteran

teachers – those with six or more years of full-time teaching experience – who also

received large raises did not express the same level of optimism about TCP as did less

experienced teachers who received the large raises. I return to the group of "young

teachers – large raises" later in the Results section.

"I believe that the Teacher Career Pathway will continue to improve over time as a resource to support teacher growth and development."		
	Jan. 2015 Survey	
Teachers on TCP	**Average**	**N**
Raise under 5% + Less than 6 yrs experience	46.57	28
Raise under 5% + More than 5 yrs experience	46.09	11
Raise over 5% + Less than 6 yrs experience	60.38	29
Raise over 5% + More than 5 yrs experience	47.81	16
0 = "not at all true." 70 = "definitely true."		

Investment in improvement process was high among Home Office staff and

school administrators, low among teachers. Responses from School Directors

indicated that they felt invested in the implementation of TCP as a learning endeavor. On

the October survey, one School Director wrote, "I feel confident in decisions that were

made as we were consistently checking the philosophical framework to make sure we

were not making Band-Aid fixes. Through all of this, I believe we really were able to recognize and celebrate teachers through the pathway." The sense of investment in an iterative, learning process was echoed by another School Director who wrote, "We made changes as we learned more about how our system was developing. And these changes reflected the founding beliefs of TCP." The Director of Human Capital pointed out, "We really strove to modify, improve, and build TCP." Use of the first person plural by all of these leaders is indicative: by the late fall, most Home Office staff and school administrators expressed collective investment in TCP's continuous improvement.

The picture was much different among teachers. While pleased to see that Home Office staff and administrators appeared to be learning, teachers expressed distance and variable investment in the improvement process. One teacher wrote, "The authors of the TCP process showed a genuine openness to improving the process in response to teacher feedback," a positive review of the feedback process that nonetheless suggested that administrators, not teachers, were seen as the "authors" of TCP. Citing a concern about the way that compensation was determined under the new system, another teacher demanded, "This needs to be fixed and not swept under the rug or called a 'learning experience for TCP,'" putting the onus for improvement on administrators.

When invited to participate in the improvement process, many teachers simply declined. Of the 99 teachers who received placements on the career pathway, only half completed the October 2014 feedback survey. Moreover, only 14 teachers participated in the four feedback sessions in November and December 2014. Among these 14 teachers, the "young teachers – large raises" group who expressed the greatest belief in TCP's ongoing evolution were no more likely to attend than were other groups, suggesting that,

while they may have been optimistic about improvement, they were not deeply invested in the improvement process themselves, seeing administrators as the primary agents of change.

Attendance at Feedback Sessions (Nov. + Dec. 2014)	
Teachers	**N**
Raise under 5% + Less than 6 yrs experience	4
Raise under 5% + More than 5 yrs experience	1
Raise over 5% + Less than 6 yrs experience	4
Raise over 5% + More than 5 yrs experience	2
No placement in 2014	3

Teachers' open responses on the January 2015 mid-year survey also bore this out. I analyzed all 125 open responses about TCP that were submitted by teachers on the survey. In each response, I looked for use of the first person plural, or related language, to describe the improvement process. Only five teachers referred to it in the first person plural – and one of these responses was from a former school administrator who had transitioned back to the classroom in 2014. Like other members of the "young teachers – large raises" group, one teacher who had rated the "belief in improvement" question highly (70 out of 70) nevertheless put the expectation and onus for improvement back on administrators, writing: "I actually think TCP has come a long way and is becoming more useful/helpful. I think the more clarity the Network can provide around TCP data measures … the better it will be." Despite believing in improvement, this teacher was not deeply invested in participating in the improvement process.

The third strand of my theory of action held that in order to shift the system towards a developmental purpose, I had to orient stakeholders around TCP as a resource for development. The results suggest variation in the orientation of stakeholders. By the end of the fall, Home Office staff and school administrators had started to orient around TCP as a developmental pathway. Meanwhile, teachers were split. Less experienced, generally younger teachers continued to view TCP as an absolute measure of teacher effectiveness, while more experienced teachers tended to understand the intent of the system as being to develop teachers at different stages along a trajectory. However, the more experienced teachers were generally pessimistic about TCP's ability to achieve its developmental purpose, while younger teachers expressed optimism about the system's ability to measure teacher effectiveness.

Home Office staff and school administrators were beginning to speak about TCP in terms that positioned it as a developmental resource. In order to evaluate the extent to which the Home Office TCP team was oriented around the system as a resource for development, I looked for evidence of a development orientation in my one-on-one interviews with team members. Specifically, I coded responses for evidence that team members saw a dual purpose for the Teacher Career Pathway as being about both (i) measurement and (ii) development. A summary is below.

Evidence from semi-structured interviews		
Team member	**Named measurement purpose**	**Named development purpose**
Director of Human Capital	X	X
Senior Manager of Data & Assessment	X	
Senior Manager of Education Technology	X	
Director of Operations	X	X
Director of Finance	X	
Director of Schools	X	X
Senior Manager of Professional Development	X	X

The data suggest some limited success at orienting the TCP team around a concept of

development. While every member of the team named one purpose of the system as

being to measure the performance of individual teachers, four out of seven explicitly

discussed the role of the system in helping teachers grow from Novice to Master. These

four team members generally pointed to a disconnect between the developmental intent

and the measurement reality: despite TCP's avowed purpose, these team members

believed that most teachers experienced the system as purely evaluative. Nevertheless,

the fact that they identified the developmental intent suggests that they recognized the

need for a development orientation. "I think we have to do a pretty significant cultural

shift to one that is focused on learning. That does not pervade a lot of mindsets.

Achievement trumps the learning in a lot of cases," said the Director of Human Capital

(J. Goldstein, October 28). The Director of Schools agreed. "It's still a metric for a lot of

people," he said, reflecting on his conversations with teachers. "[There is] not really

belief that it is a development tool…. [We] need to spend some time building shared vision" (S. McVoy, October 30). School administrators similarly recognized the gap between reality and intent of TCP. One middle school Dean wrote, "I think that [TCP] is helping to push the conversation around meaningful professional growth and meaningful recognition, but we still have a long way to go."

Teachers mainly oriented towards TCP as a high-stakes measurement tool. In order to understand how teachers thought about TCP, I analyzed open responses on the January 2015 network-wide mid-year survey, which occurred three months after the fall TCP placement. After answering six ratings questions, teachers were given the following prompt: "Please use this space to elaborate on any of your responses above." The open-ended wording of this question provided a window into the perspectives of DSST's teachers, allowing them to express an orientation to the system. As noted earlier, there were 125 responses. For this analysis, I chose to look only at the responses of teachers who had received placements on the pathway in 2014, believing that this group of 55 teachers would have the most experience and the most sophisticated understanding of the system. I coded responses with one of four values, according to the table below.

Value	Response type	Example
1	**Development**. Writer points out shortcomings or suggests improvements that would help teachers to grow professionally as a result of the Teacher Career Pathway.	"In order for this pathway to really help people chart a course for their professional development over the course of a career, opportunities that individuals have available to them at each stage in their career need to be well explained…"
2	**Measurement: weighting and pathway placement**. Writer points out shortcomings or suggests improvements to the way that measures are weighted and pathway placements are produced.	"I think we need to incorporate more of the unannounced informal observations into our data. Why do those not get counted towards our results if our appraiser is there anyways?"

| 3 | **Measurement: instruments**. Writer points out shortcomings or suggests improvements to the evaluation instruments themselves. | "First, I think all TCP assessments should be blind. Second, observations should be done by somebody who doesn't know the teachers. It could be somebody from another campus…but there are just too many biases when people observe friends and colleagues they work with daily." |
| 4 | **General/unclear.** Writer's orientation towards the system cannot be deduced from comment. | "It's not perfect; but I appreciate how teacher feedback has been incorporated over the past 3 years." |

Next, I looked for patterns in the open responses. Overwhelmingly, the responses focused on the measurement function of TCP, with 39 of 55 responses (71%) falling into the second or third categories described above. However, there was a slight difference among teachers based on their years of teaching experience: more experienced teachers tended to orient more towards a developmental purpose than did less experienced teachers. Among the open responses of 19 teachers with six or more years of teaching experience, six teachers (32%) focused on the development purpose of TCP compared to just two (5%) of the 40 teachers with less than seven years of experience. Indeed, among the open responses of teachers with less than six years of experience, 75% focused on measurement.

Years of teaching experience	Open response type				
	Development	Measurement: weighting and placement	Measurement: instruments	General /unclear	Total
1-3	2	9	7	2	20
	10%	45%	35%	10%	100%
4-5	0	8	3	5	16
	0%	50%	19%	31%	100%
6+	6	6	6	1	19
	32%	32%	32%	5%	100%
TOTAL	8	23	16	8	55
	15%	42%	29%	15%	100%

Looking solely at years of experience, the difference between teachers with 1-3 years of experience and those with 4-5 years appears slight. However, by combining years of experience with size of raise, a more nuanced picture emerges. Early-career teachers who received a standard raise (generally 0%-3%) in 2014 sought improvements to the weightings and policies that produced career pathway placements – 58% of their open responses focused on the second category described above. Meanwhile, the "young teachers – large raises" group was as equally focused on improvements to the weightings and placements (35%) as they were focused on evolution of the evaluation instruments themselves (41%). In other words, early-career teachers consistently oriented around the measurement purpose of TCP; however, those who received a standard raise (0%-3%) as a result of the system were likely to seek big, substantive changes to the weightings and placement policies.

Teachers on TCP			Open response type			
		Development	Measurement: weighting and placement	Measurement: instruments	General /unclear	Total
Raise under 5% + Less than 6 yrs experience		1	11	3	4	19
		5%	58%	16%	21%	100%
Raise under 5% + More than 5 yrs experience		2	2	3	1	8
		25%	25%	38%	13%	100%
Raise over 5% + Less than 6 yrs experience		1	6	7	3	17
		6%	35%	41%	18%	100%
Raise over 5% + More than 5 yrs experience		4	4	3	0	11
		36%	36%	27%	0%	100%
TOTAL		8	23	16	8	55
		15%	42%	29%	15%	100%

In sum, the results from the "If" side of my theory of action point to the following:

1. Between July 2014 and January 2015, I was able to cultivate a cross-functional development team that approached TCP's implementation as a learning endeavor.

2. In schools, administrators shared the view of implementation as learning. Teacher opinion was far more mixed and, across the board, teachers lacked investment in improvement, seeing administrators and Home Office staff as the primary agents of change.

3. A small group of Home Office team members, school administrators, and experienced teachers considered the possibilities for TCP to become a developmental resource, but opinion was still overwhelmingly oriented towards the system as a measurement tool.

Then: I will foster flexibility that allows DSST to begin to adapt the Teacher Career Pathway

What occurred as a result of my efforts to build a cross-functional development team, frame implementation as learning, and orient stakeholders around a concept of development? I look at two sources of evidence: (i) changes to the TCP system itself and (ii) changes in stated perceptions of the system by DSST stakeholders.

By late January 2015, four concurrent improvement processes were underway. Within weeks of completing the first cycle of TCP, significant improvement processes were underway. Each of these processes aimed to effect substantive changes to the TCP system itself. Importantly, however, all four improvement processes targeted the measurement purpose of TCP. I describe the processes briefly.

1. **Survey administration**. Seeking to provide a cleaner user interface and more reliable administration process, the Director of Operations and Senior Manager of Education Technology, with the collaboration of the TCP team, undertook a rapid

process in December 2014 to transition student and parent surveys (a key piece of the TCP measurement tool) from a homegrown system to a third-party system built and managed by Panorama Education. Unplanned as late as November 2014, this transition occurred as a result of dialogue among the TCP team and School Directors about ways to provide better student and parent feedback visualizations and more reliable data to teachers.

2. **Survey design and content**. Responding to teacher feedback, I launched a process in January 2015 to engage teachers and leaders in a redesign and rewrite of all student, parent, peer, and supervisor surveys that formed part of TCP. This process was intended to provide teachers with more actionable feedback and to eliminate bias in the survey questions. Feedback from teachers had consistently pointed to shortcomings in the survey content. This process was intended to correct those shortcomings.

3. **Rubric and observation**. Also responding to teacher feedback, the Chief of Schools and Director of Curriculum and Instruction introduced a process in January 2015 to differentiate DSST's instructional rubric and observation process by stage on the pathway. Of the four improvement processes underway in 2015, this project was most likely to serve a developmental purpose. Indeed, it was intended to better align observational feedback with the career stage of the teacher.

4. **Compensation**. The final improvement process involved a review of the TCP financial model. Led by DSST's Chief Operating Officer and Director of Human Capital, this project surfaced fundamental questions: what was the best way to

align compensation with movement on the career pathway? How should

compensation be structured to incentivize long-term retention and performance?

Taken together, these projects pointed to a willingness and flexibility to improve the

Teacher Career Pathway, but they were oriented more towards a concept of TCP as a

measurement tool than they were geared towards TCP as a developmental resource.

Indeed, while improvements to the surveys and observation process contained a

developmental purpose – better data could be used to pinpoint areas of growth – these

changes were largely a response to teachers who voiced an interest in improving

measurement tools in order to generate more reliable indicators of teacher effectiveness.

A project to closely align placement on the pathway with professional development

opportunities and instructional coaching was slow to take off.

Changes in perception tracked teachers' experiences with the system. In

order to capture changes in teachers' perception of TCP, I ran a matched comparison of

responses on a May 2014 End-of-Year Survey and January 2015 Mid-Year Survey. Of

the 99 teachers who received placements in Fall 2014, 84 responded to both surveys,

allowing me to produce matched comparisons. I looked at answers to all TCP-related

questions, but most closely examined responses to "In charting my career pathway, the

Teacher Career Pathway is helpful." I believed that evaluating changes in responses to

this question would be a reasonable proxy for overall changes in perception of the

system. I found that changes in perception tracked teachers' experiences with TCP.

Teachers who initially "placed" in the third, fourth, and fifth stages on the pathway rated

the question more highly and showed greater improvements in perception than did

teachers who placed in the first and second stages. This trend was most pronounced among teachers at the Lead Teacher stage.

"In charting my career pathway, the Teacher Career Pathway is helpful."				
TCP Stage	EOY 13-14	MY 14-15	Change	N
Novice	40.38	37.63	-2.75	8
Developing	32.74	33.17	0.43	23
Accomplished	40.07	45.9	5.83	29
Lead	39.38	49.13	9.75	16
Master	41.88	43.5	1.62	8
Average	37.5	39.66	2.16	84

Moreover, improvements in perception largely followed the size of the raise that teachers received as a result of TCP. Many of the largest raises were concentrated in the Accomplished and Lead stages, since Master Teachers tended to be more experienced and have higher salaries under the old salary scale.

"In charting my career pathway, the Teacher Career Pathway is helpful."				
Raise	EOY 13-14	MY14-15	Change	N
No raise	29.60	29.10	-0.50	10
2%-4.99%	34.21	35.97	1.76	29
5%-10.99%	40.44	45.78	5.34	9
11%-20%	43.82	49.73	5.91	22
Over 20%	41.92	49.21	7.29	14
Average	37.5	39.66	2.16	84

In other words, "doing well" on the career pathway tended to correlate with both higher overall perception of the pathway and with greater improvements between May 2014 and January 2015. Again, however, bringing years of experience together with size of raise produced a more nuanced picture. Compared to other groups of teachers, the "young teachers – large raises" group rated this question about the helpfulness of TCP 14 to 19

points higher overall and showed the greatest numerical improvement (11.17) on the

January 2015 survey compared to the May 2014 survey.

"In charting my career pathway, the Teacher Career Pathway is helpful."				
	EOY13-14	MY14-15	Change	N
Raise under 5% + Less than 6 yrs experience	34.56	34.70	0.15	26
Raise under 5% + More than 5 yrs experience	28.64	34.27	5.64	9
Raise over 5% + Less than 6 yrs experience	41.93	53.10	11.17	29
Raise over 5% + More than 5 yrs experience	44.06	39.24	-4.82	17
Average	37.5	39.66	2.16	84

Simply put, 46 teachers (in the matched comparison) received raises over 5%, but

only those teachers with less than six years of teaching experience showed improvement

in perception of TCP. Perception actually *declined* by 4.82 among the 17 teachers who

received raises over 5% and had more than five years of teaching experience. This

pattern was true on other questions as well: the "young teachers – large raises" group had

higher overall ratings of TCP and showed greater improvement in perception compared

to other groups.

"The Teacher Career Pathway surveys (peer, student, supervisor, and advisory) provide me with fair and informative feedback."				
	EOY13-14	MY14-15	Change	N
Raise under 5% + Less than 6 yrs experience	41.56	41.44	-0.11	26
Raise under 5% + More than 5 yrs experience	37.36	29.45	-7.91	9
Raise over 5% + Less than 6 yrs experience	46.48	52.10	5.62	29
Raise over 5% + More than 5 yrs experience	47.94	43.47	-4.47	17
Average	42.8	45.78	2.98	84

"The observational process provides me with fair and informative feedback."				
	EOY13-14	MY14-15	Change	N
Raise under 5% + Less than 6 yrs experience	42.30	47.52	5.22	26
Raise under 5% + More than 5 yrs experience	36.73	39.73	3.00	9
Raise over 5% + Less than 6 yrs experience	50.59	54.55	3.97	29
Raise over 5% + More than 5 yrs experience	44.82	45.71	0.88	17
Average	44.41	49.13	4.72	84

Overall, among teachers on the pathway, perception of TCP increased slightly between May 2014 and January 2015. However, one identifiable group accounted for much of the improvement in perception. Results introduced earlier in this section indicate that this group was optimistic about TCP's continuing evolution, but was not oriented towards TCP as a developmental resource and was no more invested in participating in improvement of the system than were the other groups of teachers.

The remainder of the capstone analyzes these results and offers implications for self, site, and sector.

3. Analysis

In this section, I combine frameworks introduced in the RKA with Williams's (2010) "real leadership" framework to examine my strategic project through a series of analytic lenses. Derived from Heifetz's theory of adaptive leadership, *real leadership* represents a useful lens because Williams, a colleague of Heifetz, has contextualized the theory for several types of "leadership challenges," including the "development challenge," a scenario where a "new opportunity opens up" through which a group "can make significant improvement to its quality of life or organizational performance if latent abilities become effective" (Williams, 2010, p. 154, 59). DSST Public Schools viewed passage of Senate Bill 10-191 as just such an opportunity – a chance to create a career pathway that would "identify, develop, and retain effective teachers," leading to significant improvements in organizational performance (DSST Public Schools, 2014). My key leadership task was "to orchestrate a learning process" that would allow the group to realize this opportunity (Williams, 2010, p. 151).

According to Williams, there are four "real leadership" strategies for a development challenge (2010, p. 164):

1. Create a robust holding environment to keep people from getting distracted.

2. Develop in stages: give the people time to discover what works.

3. Find the right combination of levers to develop new values and capabilities.

4. Give people a stake in developing their capacity.

In the following analysis, I argue that a variety of factors, including my own leadership moves, created a robust holding environment for the TCP team and for School Directors. That holding environment, coupled with framing that presented TCP's implementation as a learning endeavor, fostered flexibility, resulting in a willingness to adapt TCP. In regards to TCP's implementation, fostering flexibility meant giving people time to identify and discuss underlying principles, check progress against those principles, and continuously modify the TCP system based on new understanding. However, two factors hindered my ability to fully position TCP as a resource for teacher growth and development. First, I engaged teachers only superficially in the learning process, failing to invest them in developing "new values and capabilities." Teachers heard the message that TCP's implementation was a "learning endeavor," but they were not genuinely invested in the endeavor. Second, teachers and administrators had a fundamentally different stake in developing TCP. For Home Office staff and school administrators, TCP could be built iteratively and dispassionately into a system to identify, develop, and retain effective teachers. For teachers, on the other hand, the system determined compensation and generated feelings of self-worth, for better or for

worse. For teachers, it was understandably difficult – and perhaps impossible – to see

beyond the shadow of measurement. I elaborate on these points in the following analysis.

Successfully fostering flexibility I: robust holding environment kept people from getting distracted. Borrowing a term from the field of psychoanalysis, Williams (2010) argues that a robust holding environment "serves to 'hold' people through the difficult and disorienting work of developing new capabilities and adjusting to new contexts" (p. 165).[3] A leader's job is to define "the features of the holding environment" (p. 165). As the leader of TCP's initial implementation, I attempted to foster flexibility that would allow DSST to begin to adapt the system towards a developmental purpose. As the results indicate, I was successful at fostering flexibility. Team members exhibited a willingness to learn from work in progress and make modifications to the system. However, the organization was slow to use that flexibility to adapt the system specifically towards a developmental purpose. The features of the robust holding environment help to explain why I was successful at fostering flexibility (and less successful at initiating developmental adaptations of the system).

Before describing the features of the holding environment, it is worth revisiting my role in TCP's implementation. Earlier, I introduced the concept of heavyweight project managers, individuals who may lack formal authority over members of a cross-functional team, but nevertheless "exercise strong direct and indirect influence across all functions and activities in the project" (Clark and Fujimoto, 1991, p. 255). Heavyweight

---

[3] Used frequently in the literature on adaptive leadership, the term "holding environment" is defined by Heifetz (1994): "A holding environment consists of any relationship in which one party has the power to hold the attention of another party and facilitate adaptive work" (p. 105). The term originally comes from the field of psychoanalysis and the work of British pediatrician Donald Winnicott (see Heifetz, 1994, p. 310).

project managers exercise strong influence by owning the internal integration of a development project and serve as champions for the concept itself – whether the concept is a product, like a new automobile, or a process, like a teacher career pathway. "The heavyweight product manager is essentially a traveling preacher," write Clark and Fujimoto, "whose bible is the formal concept and planning documents" (1991, p. 262).

From July 2014 to January 2015, I acted as what I would term a "middleweight project manager." I owned the internal integration of the project, overseeing and coordinating the disparate functions of academics, operations, finance, education technology, data, and human capital. However, I was a champion not for TCP as it was defined in the TCP Guidebook, the bible produced over three years by the original Working Group, but rather for a concept of TCP's implementation as a learning endeavor that would be open to modifying, adapting, and potentially tossing out elements of the bible. Whereas a heavyweight project manager would have advocated for the original concept, I recognized that the concept was not serving its purpose of supporting teacher development and tried instead to create a climate that would result in adaptation of the bible. Indeed, I neither championed the original TCP Guidebook nor introduced a new version of the guidebook myself; instead, I tried to create an environment that would 'hold' others through the production of a renewed concept and, I hoped, a new bible. The figure below, adapted from Clark and Fujimoto (1991), aims to demonstrate how I created a holding environment around the TCP team.
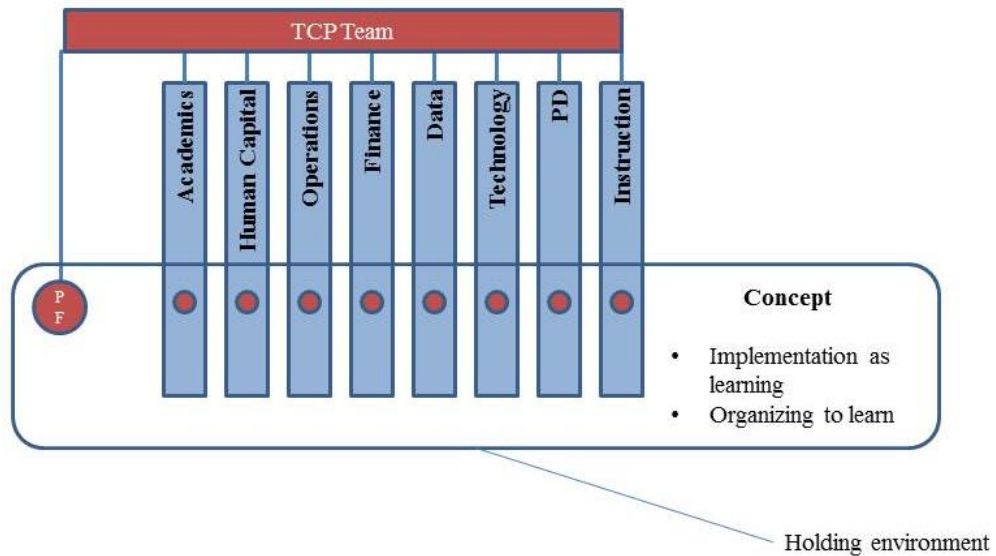
Edmondson's (2012) "organizing-to-learn" framework defined the features of this holding environment. By establishing guiding principles, modeling an inquiry orientation, and encouraging dialogue and double-loop problem-solving, I sought to create a climate on the Home Office's TCP team that was nimble and open to iteration. As evidenced by the willingness of TCP team members to launch substantive improvement processes immediately after completion of TCP's initial implementation and to iterate on the system even in the midst of that initial implementation, the holding environment, defined by "organizing-to-learn" principles, proved effective at fostering flexibility.

However, simply defining the features of an environment is not enough to 'hold' people or to keep them from getting distracted. Several factors proved critical to establishing the legitimacy of "organizing-to-learn" principles as the defining features of the holding environment.

Securing a quick win. At the first meeting of the TCP team in July 2014, we had agreed to "work collaboratively to understand individual responsibilities as well as their

fit within the larger system." Within days, this commitment was put to the test. In 2013, without consulting other members of the Home Office team, the former Director of Human Capital had set September 30, 2014, as the deadline for completing TCP's initial implementation. During a TCP team dialogue in July, I guided the group to the discovery that this timeline was not feasible for a variety of reasons. Working together, the group identified an underlying cause (lack of open, regular communication among the team), created an integrated plan for pushing back the implementation timeline, consulted with School Directors and a teacher focus group, and finalized the plan within ten days of surfacing the problem. The Senior Manager of Data and Assessment said, "[Before Summer 2014] I was never part of a decision to do this by the end of September. I don't remember weighing in. My attitude was that I'll execute what is needed" (S. Coit, October 28). For members of the newly constituted TCP team, the process demonstrated the value of collaboration, interdependent thinking, and psychological safety – the type of "new models of behavior" that I hoped to instill in the group (Watkins, 2009). This "quick win" built buy-in to the features of the holding environment.

Modeling vulnerability. A second factor that established the legitimacy of "organizing-to-learn" was my own effort to model vulnerability. In August, the team received word that the results of statewide testing, a component of the TCP evaluation, would not be released for the 2014-15 school year until as late as Spring 2016, rendering them inoperable for the 2014-15 TCP cycle. An alternative was needed. As a TCP team, we developed an alternative, incorporating input from members of the Home Office's Academic Team through a series of conversations and meetings over the course of one week, communicating an alternative to teachers on one of the final days before the school

year began. The process was not as inclusive as many administrators and teachers hoped, leading to several complaints that diverse perspectives had not been included in the process. These complaints were justified. In trying to move quickly and communicate before the start of the school year, we had missed opportunities to engage a broad group of stakeholders. With the Director of Curriculum and Instruction and Director of Human Capital, I told the TCP team and School Directors by email that we had made a mistake, moving too quickly and not including enough voices in the process. We offered support to School Directors in addressing our own misstep. Very publically, I, along with other Home Office leaders, modeled vulnerability. Coming on the heels of an inclusive process to push back the fall's implementation timeline, this incident affirmed a climate of psychological safety, where members of the team were free to make mistakes without fear of penalty. After all, several key leaders had made a mistake and owned the consequences.

   Upholding principles in the face of a high-stakes deadline. A final factor that helped to form the holding environment was the existence of a high-stakes deadline – and the team's ability to avoid "the performance pressure paradox," the tendency of groups to shut down communication, defer to authority, and opt for safe solutions when faced with a deadline (Gardner, 2012). Gardner argues that high-pressure situations often lead to "fewer people … at the usual meetings" and team members "literally turning their backs on colleagues offering alternative views" (2012, p. 85). At several points between August and October 2014, members of the senior leadership team sought to intervene when errors arose from the TCP team. Rather than deferring to authority and shutting down communication, I maintained the primacy of the team and the importance of

81

collaborative, double-loop problem solving. This did not go unnoticed. The Director of Human Capital would later say to me, "The outcome can frequently trump the team, but we had a commitment to this team finishing all of this work" (J. Goldstein, October 28).
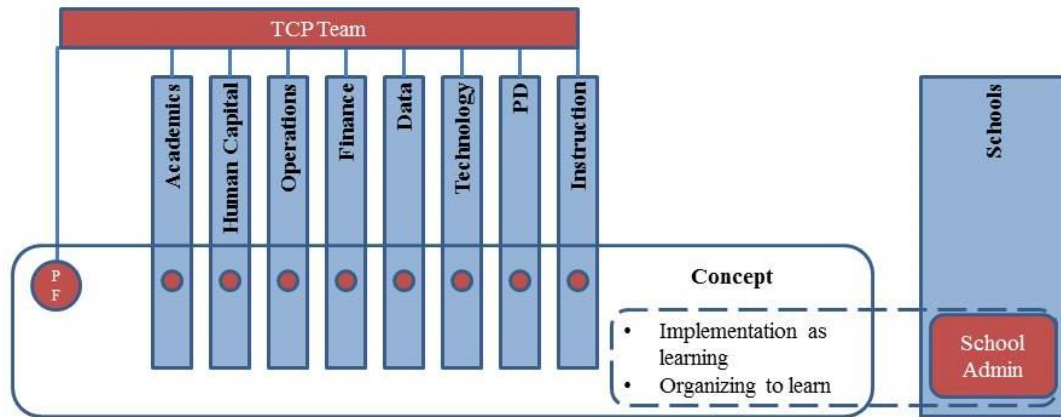
These three factors – securing a quick win, modeling vulnerability, and upholding teaming principles in the face of a high-stakes deadline – helped to establish "organizing-to-learn" as the features of a holding environment that kept attention on TCP's implementation as a learning endeavor and kept pressure on the team to act nimbly and flexibly. The TCP team could have gone into what Edmondson (2008) calls a mode of "execution-as-efficiency," where subordinates look to leaders to provide answers and teams fail to engage in authentic dialogue and collaborative problem-solving. Instead, I took an "empower rather than control" approach with the TCP team, championing a concept of TCP as a continuously evolving system that would improve over time as a resource to guide teacher growth and development (Edmondson, 2008, p. 67). By empowering rather than controlling, I helped to "contain and regulate the stresses" that learning generated on the TCP team (Heifetz, 1994, p. 105).

Successfully fostering flexibility II: gave (some) people time to discover what works. "Getting it right the first time is rarely possible in a development challenge," writes Williams (2010, p. 178). "It is a process rife with uncertainty, breakdowns, and, sometimes, quantum leaps of understanding and progress" (2010, p. 174). To lead groups through a development challenge, the leader must give people time to work through uncertainty, learn, and gain deeper understanding. He or she must "manage processes that get the people to do the sometimes tedious work of discovering for

themselves which practices and priorities actually lead to increased prosperity for the community, the organization, or the nation" (2010, p. 182).

On the TCP team, time was devoted to dialogue, which occupied dozens of hours between July and October. By devoting weekly team meetings to dialogue, rather than simple report-outs or presentations, I kept the team's attention on collaboration and learning. As discussed in the Results section, team members recognized in interviews that the amount of time dedicated to dialogue was critical to building shared understanding and belief in flexibility as a guiding idea behind TCP's implementation.

Time spent in dialogue also proved critical for investing School Directors in the concept of TCP's implementation as a learning endeavor. By spending at least half (3-4 hours) of each monthly, network-wide Leadership Team meeting devoted to TCP's implementation between July and October, I was able to extend the holding environment that governed the TCP team out to School Directors, enlisting them in the learning endeavor. During those meetings and through one-on-one coaching and communication with me and the Director of Human Capital, School Directors experienced and became invested in the iterative, experimental nature of implementation, as evidenced not only by their ratings on the October 2014 and January 2015 surveys, but also by comments like that of the School Director quoted earlier who said, "I feel confident in decisions that were made as we were consistently checking the philosophical framework to make sure we were not making Band-Aid fixes." As the figure below indicates, School Directors gradually became invested in implementation as learning.

TCP Team

Academics · Human Capital · Operations · Finance · Data · Technology · PD · Instruction

Schools

P F

Concept
- Implementation as learning
- Organizing to learn

School Admin

However, lack of time also proved a hindrance to deeper investment. While School Directors engaged in dialogue during monthly meetings and one-on-one communication with members of the TCP team, other members of school administrative teams – especially Deans and Associate School Directors – were far less involved in the learning process. Moreover, after the initial implementation of TCP was completed in October, time devoted to dialogue and learning among School Directors quickly gave way to other concerns. For instance, at December's network-wide Leadership Team meeting, a session that I had designed to enlist School Directors in constructing a shared vision for development was replaced by work with the team to address short-term teacher retention concerns. The daily exigencies of running schools loosened the holding environment around school administrative teams, resulting in improvement processes that were driven mainly by the Home Office. For school administrative teams, "immediate responsibilities supersede[d] potential future benefits of improvement efforts" (Tucker, Edmondson, Spear, 2002, p. 133).

Failure to produce positioning of TCP as a developmental resource I: used the wrong combination of levers. By January 2015, the TCP team and, to a somewhat
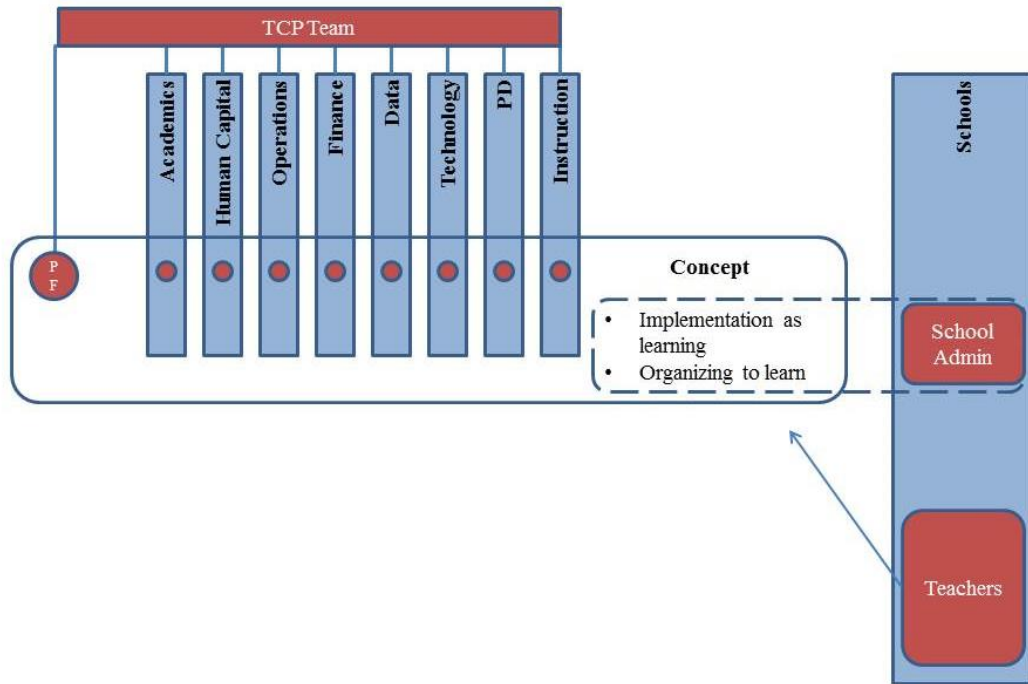
lesser extent, school administrators exhibited flexibility and a willingness to adapt the

system. However, why was that flexibility directed primarily at improving the ability of

the system to measure performance, not at building the capacity of the TCP system to

develop teachers? I argue that there were two crucial reasons. First, although teachers

witnessed the TCP team and school administrators engaged in a learning endeavor, I

chose the wrong combination of levers to invest teachers in that endeavor.

Simply put, I used what Chris Argyris called communication tools that "block

learning" (1994, p. 77). From August through January, I invited teacher participation

through three distinct surveys, four feedback sessions, and a continuous invitation to

submit feedback both through school administrative teams and directly to the Home

Office. These communication tools resulted in a tremendous amount of data on teacher

perceptions of TCP. However, they did not invest teachers in the concept of TCP's

implementation as a learning endeavor. As Argyris writes, surveys and feedback sessions

"produce useful information about routine issues like cafeteria service and parking

privileges," but:

> What they do *not* do is get people to reflect on their work and behavior.
> They do not encourage individual accountability. And they do not surface
> the kinds of deep and potentially threatening or embarrassing information
> that can motivate learning and produce real change (1994, p. 77-78).

Rather than looking inward and examining their own responses to and engagement with

TCP's implementation, most teachers placed the onus for improvement elsewhere, saying

things like, "This needs to be fixed and not swept under the rug or called a 'learning

experience for TCP.'" Teachers' reactions should have been unsurprising: requests for

feedback asked for input, but did not enlist teachers in self-reflection or learning. The

result was that teachers remained detached from the entire endeavor of TCP's

implementation and ongoing improvement, as illustrated by the diagram below.

TCP Team

Academics | Human Capital | Operations | Finance | Data | Technology | PD | Instruction

PF

Concept

- Implementation as learning
- Organizing to learn

Schools

School Admin

Teachers

Why did I choose communication tools that "blocked learning"? One reason is

my own defensive routines. Engaging teachers in deep self-reflection of their experience

with and responses to the Teacher Career Pathway would have surfaced what Roland

Barth (2006) calls "non-discussable" topics, such as questions about the very premise of

using quantitative measures to evaluate teacher performance and the tendency of DSST's

mostly young team of teachers to be "subject" to these measures. In my work on another

project at DSST, I had received criticism from School Directors and the Chief of Schools

for surfacing non-discussable topics during a series of sessions with school teams in

September and October. In an effort to remain in control and avoid further unpleasant

feedback, I believe that I implicitly chose to engage teachers only superficially in the learning process around TCP.

Moreover, it was easier to demonstrate responsiveness to feedback from teachers in the form of surveys and focus groups than it would have been to demonstrate responsiveness to genuine self-reflection. As Argyris points out, defensive routines include the tendency to be "as rational as possible, by which we mean laying out clear-cut goals and then evaluating our own behavior on the basis of whether or not we've achieved them" (1994, p. 80). In the form of surveys and focus groups, feedback from teachers centered on "technical" improvements to TCP, including changes to the content of student course surveys and tweaks to the weighting of evaluative components. Compared to the adaptive work of transforming TCP into a developmental resource, technical changes focused primarily on TCP's measurement function could be easily explained as a set of clear-cut goals (e.g. rewrite student course survey content). Progress could be easily demonstrated against these goals. As discussed in the Results section, technical changes to TCP's measurement function became the centerpiece of improvement efforts in late 2014.

Failure to invest teachers in TCP's implementation as a learning endeavor stemmed not only from my own defensive routines, but also from those of the organization. Two points about DSST are worth repeating. First, as a K-12 system of schools, DSST is part of an industry that Senge described as "the starkest example in modern society of an entire institution modeled after the assembly line" (2000, p. 30). In a bureaucratic system, administrators control the teachers and teachers control the students. From the classroom to the Home Office, DSST team members tried

extraordinarily hard to avoid bureaucratic theory, institutionalizing practices and rituals that aimed to soften the divisions between administrators and teachers. However, the system remained structured as a top-down bureaucracy – with decisions at the Home Office flowing down to impact school administrators, who subsequently made decisions that served to control teachers. Surveys and focus groups buttressed this structure. As Argyris notes, "By assigning all the responsibility for fixing problems to management, [surveys and focus groups] encourage managers *not* to relinquish the top-down, command-and-control mindset that prevents empowerment" (1994, p. 83).

Second, as a charter school organization in fierce competition for students and for funding, DSST exhibited many of the characteristics of the market-oriented culture described by Cameron and Quinn (1999), particularly an intense task and achievement orientation. Administrators tracked progress on a wide range of organizational indicators, including teachers' perceptions of TCP. Staff survey data were used for both formative and summative purposes – to pinpoint areas for improvement and to measure the effectiveness of the organization and of administrators. This culture created incentives to seek the type of feedback from teachers that could be easily addressed and where the organization's ability to address feedback could be easily measured. Just as I wanted to be as rational as possible, so too did the organization aim to rationalize teachers' experience with TCP, distilling feedback down to technical changes that could be executed rather than engaging teachers in the messy work of adaptive change.

Failure to produce positioning of TCP as a developmental resource II: people had a fundamentally different stake in developing TCP. In order to help the group build new capabilities, leaders must give people a stake in developing new capabilities.

With passage of SB 10-191, DSST Public Schools faced an opportunity to build a new capability, namely the ability to systematically identify, develop, and retain great teachers. Unlike in many other development challenges, all members of the team had a stake in developing this capacity. Indeed, giving people a stake was not the problem. The problem was that Home Office and school administrators had a fundamentally different stake in TCP than did teachers. This difference helps to explain much of the variation in perceptions of TCP as well as my inability to position TCP as a developmental resource.

For DSST's Home Office and school administrators, the promise of TCP was in creating a system that would identify the effectiveness of teachers, create opportunities for recognition and reward that would retain the most effective teachers, and highlight areas where administrators could better support less skillful educators – all with the goal of improving teacher quality and thereby increasing student achievement. A career pathway for teachers meant better results for children. Administrators could speak dispassionately about improvements to TCP and approach development of this organizational capability as a learning endeavor because they were not directly and personally implicated in the system. In the language of adult development, administrators could stand outside of the TCP system, taking it as "object."

The stakes were far different for teachers. By attaching labels (e.g. "Novice," "Developing," etc.) to teachers and determining their compensation, the Teacher Career Pathway acutely impacted teachers' feelings of self-worth, not to mention their livelihoods. With such direct and personal stakes attached to development of this organizational capability, teachers justifiably struggled to stand outside of the system,

taking it as object, and instead felt subject to its ongoing evolution. Thus, rather than viewing TCP's iterative implementation as a signal of learning, teachers feared that an evolving system would unfairly label them. As one teacher wrote on the January 2015 network-wide survey, "I have a problem when an unfinished evaluation system determines your salary." These fears were likely compounded by the fact that surveys and feedback sessions reinforced the idea that responsibility for TCP's evolution lay with administrators, not teachers, as discussed earlier.

Given the personal stakes attached to TCP's evolution, most teachers homed in on measurement, not development. Their paramount concern was that the TCP system reliably measured their performance. Indeed, when given an open response opportunity on the January 2015 survey, 71% of responses focused on the measurement function of TCP. During the November feedback sessions, I noted twenty-four specific, suggested improvements to the system. Of these twenty-four, only three distinctly focused on adapting the system to be a developmental resource. The majority zeroed in on improving measurement. Again, this finding should be unsurprising. Teachers had experienced the system as a measurement tool; they had yet to experience it as a developmental resource.

The young age, inexperience, and achievement orientation of many DSST teachers may also have bolstered the tendency to view TCP exclusively as a measurement tool. Although I did not collect data on the stages of adult development of DSST teachers, the fact that a majority (69%) are between the ages of 21 and 30 indicate that many teachers are likely at the "socialized" stage of development. At this stage of development, one is acutely "shaped by the definitions and expectations of [one's]

personal environment," according to Kegan and Lahey (2009, p. 17). Viewing TCP

exclusively as a measurement tool, younger, less experienced DSST teachers were

pleased with the system and optimistic about its continuing evolution if they performed

well on the tool. They were far more negative and pessimistic about the future if they

received a lower placement on the pathway. Thus, teachers who received a raise over 5%

and had less than six years of teaching experience expressed consistently positive reviews

of TCP. For them, the measurement function worked well. Still, viewing TCP

exclusively as a measurement tool, this group continued to register feedback that focused

on improving measurement, not development.

More experienced teachers tended to see some slight possibilities for TCP to

function as a development tool. Indeed, among the small number (8) of open responses

devoted to development, six were from teachers with more than five years of experience.

Moreover, the group of more experienced teachers who did very well as a result of TCP

(those with more than five years of experience who received a raise above 5%) did not

exhibit the same high levels of approval for TCP, suggesting that this group may not have

been as significantly influenced by TCP's measurement and might instead have stepped

back "enough from the social environment to generate an internal 'seat of judgment'"

about the system (Kegan & Lahey, 2009, p. 17).

Despite this limited evidence from more experienced teachers, DSST's teachers

overwhelmingly oriented around TCP as a measurement tool, even as Home Office and

school administrators espoused possibilities for TCP to be both a measurement and a

development resource. "There is dissonance in the message that this is a development

tool," said the Director of Operations. "[Teachers] see it as really evaluative.… If I'm a

skeptic, this is a pretty message around a heavy evaluation system" (J. Roddy, October 30). The fact that administrators and teachers viewed TCP differently should not be surprising. The two groups had fundamentally different stakes in the system's development.
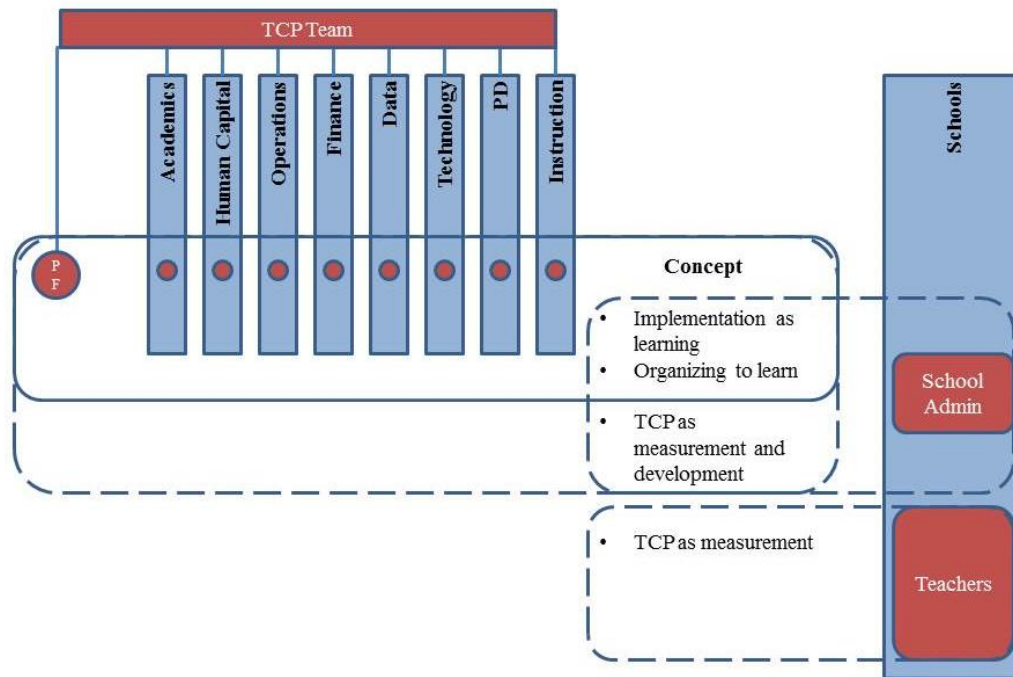
Could things have been different? The fraught coupling of measurement and development. The subtly contrasting views of administrators and teachers raise a more fundamental question: is it possible to effectively couple measurement and development in a single evaluation system? The research community is split on this question. Papay (2012) argues that "there is mounting evidence that rigorous, standards-based evaluations can improve teacher effectiveness," but Papay points to just a single study, by Taylor and Tyler (2011), which found that the performance of students in classrooms of Cincinnati Public Schools elementary and middle school teachers increased in the year that teachers

were evaluated and that these effects persisted in the years following evaluation (Papay, 2012, p. 134; Taylor & Tyler, 2011; Taylor & Tyler, 2012). Papay asserts that positioning evaluation systems as development tools is simply a matter of "refocusing the debate" away from measurement and towards development by framing evaluation data as "specific and meaningful feedback" (2012, p. 137, 138).

Meanwhile, in recent work surveying teacher development efforts across the 20th and 21st centuries, Elizabeth Green notes that "the early signs [are] not promising" that evaluation systems can improve professional learning (Green, 2014, p. 305). Indeed, Hill and Grossman (2013) contend that Taylor and Tyler's Cincinnati study is "an exception to [the] rule" and that scant evidence exists to support the claim that evaluation systems on their own improve teacher effectiveness (p. 372). Hill and Grossman argue that in order for evaluation to serve as a teacher development tool, evaluation data, primarily via classroom observations, must be subject-specific and coupled with intensive coaching and support of content experts. Popham (2013) goes a step further, maintaining that the mere fact that evaluation and evaluators affect a teacher's livelihood means that teachers will not use evaluation as an opportunity to engage in the type of candid self-reflection and learning necessary to improve practice. Evaluation systems should try to measure or develop, Popham contends; they should not try to do both.

This capstone did not set out to study whether an evaluation system could improve teacher effectiveness, but the research base surrounding this question shines light on the results obtained by the strategic project. DSST lacked a subject-specific teaching rubric and the intensive coaching and support of content experts, the areas identified by Hill and Grossman as being critical for aligning evaluation and

development. Instead, DSST used a content-agnostic instructional rubric and coaching occurred, sometimes intermittently, via school administrators who often lacked expertise in the subject. (In DSST's model, coaching included not only classroom observation and feedback, but also joint analysis of student achievement data and student, parent, and peer survey results). Moreover, administrators often served as both formal evaluator and coach – a complicating factor, according to Popham. In the absence of a coherent and aligned system of coaching to translate evaluation results into specific and meaningful feedback, it is not unexpected that DSST's teachers overwhelmingly oriented around TCP as a measurement tool, not a development tool. As one teacher said during a feedback session: "It's hard to transform our thinking from good, better, best. And you want to be the best. It's going to take a lot more intentionality in terms of introducing it as a journey, a path. I want to stay in the classroom because that's my passion, but I don't see a difference between what I'm doing now and where I want to go as a teacher."

Indeed, perhaps the most fundamental reason that teachers did not orient around TCP as a development tool is because they had experienced it only as measurement. While Home Office staff and school administrators could glimpse beyond the present state to a time when instructional coaching and professional development sessions aligned to the measurement function of TCP, these opportunities had not yet been defined and therefore were not yet visible to teachers, who understandably seemed to base their estimation of the system on their present, lived experience. That experience was one of measurement.

III. Implications for Self, Site, and Sector

By the end of January 2015, my strategic project had fostered flexibility and a willingness to adapt the Teacher Career Pathway, but the TCP system was not yet positioned as a resource for growth and development. In the prior section, I analyzed reasons for this disconnect. In the following section, I explore implications for myself as an educational leader, for DSST as an expanding system of public schools, and for the K-12 sector as a whole.

1. Implications for Self

Leading a project of both initial implementation and concurrent evolution and improvement required me to use a range of leadership tools, including the ability to diagnose context, frame implementation as learning, establish psychological safety, balance advocacy and inquiry, and facilitate communication across diverse groups of stakeholders. As I argued in the prior section, these leadership moves proved effective at fostering flexibility. What they did not do – what I failed to do – was to set a vision not only for a learning process, but also for the substantive product, meaning the next phase of TCP's evolution into a development tool, and to motivate others to align their energies behind the vision. I unpack this observation below.

Know when to employ tempered radicalism – and when not to. Entering DSST as an Ed.L.D. resident who lacked historical knowledge, credibility, and positional authority, I tried to adopt the stance of a *tempered radical*, whom Debra Meyerson describes as leaders "dedicated to their companies and masters at changing organizations at the grassroots level" (2001, p. 93-94). Tempered radicals, Meyerson writes, "work

quietly to challenge prevailing wisdom and gently provoke their organizational cultures to adapt…. In so doing, they exercise a form of leadership within organizations that is more localized, more diffuse, more modest, and less visible than traditional forms – yet no less significant" (2001, p. 93).

I chose tempered radicalism for two reasons. First, I believed that the context required it. I entered an organization of incredibly smart, passionate, and talented people working hard to build a system of schools dedicated to closing achievement gaps and ending educational inequity. I was – and continue to be – amazed and inspired by my colleagues. When I began the residency in July 2014, I lacked the knowledge, credibility, and authority to lead out in front and believed strongly that my colleagues, not I, had answers to the organization's most pressing challenges. Add the complication of having the word "Harvard" attached to my name and I deliberately decided to lead quietly, modestly, and incrementally.

Second, tempered radicalism comes naturally to me. It's my default position. I've recognized this tendency in myself for years and numerous 360 reviews have affirmed it. I approach change in one, three, and five-year increments, and try to build capacity and shift mindsets for evolutionary, not discontinuous, change. I moved easily in this position during my residency, which did not go unnoticed. As one colleague wrote on a mid-year peer review, "You have pushed to innovate our work in TCP, but have done it in a way that brings along our team and that ensures folks don't feel the boat rocks too far."

As a leadership stance, tempered radicalism allowed me to establish team processes and a holding environment centered on "organizing-to-learn" principles of

inquiry, psychological safety, and collaboration. Early on in my residency, this stance helped me to build credibility and a foundation for trust on the TCP team. The willingness to evolve TCP rested on this foundation, as team members, trusting in one another and approaching implementation as learning, proved comfortable modifying the system even in the midst of its initial implementation.

However, after the completion of initial implementation in October 2014, I continued to use the tools of tempered radicalism – gently provoking the organization and bringing people along through team dialogue and collaboration – even though the context and the task had shifted from implementation to improvement. While I was able to build capacity and shift mindsets for evolutionary change, I assumed (wrongly, it turned out) that by creating a climate of collaboration, the TCP team would craft a shared vision that effectively coupled the measurement and development purposes of teacher evaluation. Instead, the "improvement" processes that emerged out of team dialogues and review of teacher feedback in late 2014 and early 2015 continued to focus on improving the ability of the system to measure performance, not to develop teachers. As the leader of the TCP project, I could have transitioned into a more assertive style of leadership, articulating a proactive vision for coaching and professional development aligned to the measurement system. However, I continued to play the role of tempered radical.

By the end of January 2015, we had built a car with an impressive dashboard – it could tell you everything that was going on when you revved the engine – but the axels did not turn and the car did not move. As explained earlier, Clark and Fujimoto (1991) describe the heavyweight project manager as a traveling salesman whose "bible" is the planning document. I thought that, if given the right conditions, the group would craft a

vision to turn the axels and move the car – that coupled the measurement dashboard with teacher coaching and development opportunities.   However, the group really needed a leader to articulate that vision. I could not rely on others to discover it for themselves. This was a fundamental mistake.

Going forward, I need to practice greater situational awareness in order to identify both the moments when tempered radicalism will enable me to help a group achieve desired results and the moments that call for more assertive, out-front leadership. When times call for a bold vision about where the group is going and the job to be done, I need to set the vision and communicate the job-to-be-done. Leading with a vision about where the group is going may come more naturally to others than it does to me, but I can't let my natural tendencies become an excuse for not taking risks and leading out front. Indeed, tempered radicalism may be my preferred stance, but building the dexterity to use multiple styles and the awareness to know when to use what will enable me to be a more effective leader.

Set the vision and then empower people to take responsibility. Simply communicating a vision is not enough to invest others. In the case of TCP's initial implementation, teachers were not invested in improvement and they oriented around the system as measurement, not development, both because we lacked a robust vision for linking measurement and development and because teachers were not empowered to own the evaluation system as a resource that would help them to get better. As will be discussed later, separating the measurement and development purposes of TCP – saying to teachers, ostensibly, TCP is a measure of your current progress and should be used for goal-setting, but you and a coach are responsible for using the data to drive your own

growth – would have empowered others and communicated that teacher and coach were accountable for the teacher's own development. Instead, through muddled messaging and communication tools that blocked learning, I continued to perpetuate the idea that TCP was done by administrators to teachers and that the evaluation measures reflected not a resource to be owned by teachers, but an indicator of a teacher's self-worth.

   Going forward, I need to take a page from Jan Carlzon, the former leader of SAS Airlines and subject of a famous Harvard Business School case. Carlzon turned around SAS by communicating a vision in clear, concise terms – "to become the best airline in the world for the frequent business traveler" – developing a straightforward strategy, and then pushing responsibility to the front lines (Carlzon, 1995, p. 6). "We couldn't order our employees to do things differently," Carlzon wrote. "Instead, we had to convey our vision of the company and convince them that they could and should take responsibility for carrying out that vision" (1995, p. 19). For years, I've resisted this notion of leadership, seeing it as what Williams (2010) would describe as "counterfeit leadership," a "great man" view preoccupied with dominance and the conviction that others should follow the resolute leader who seems to have all the answers (p. 48-49). While I've always recognized the importance of this approach in times of crisis or other short-term problem solving, I've never believed that it builds capacity for sustained organizational change.

   The results of my strategic project have caused me to reconsider these assumptions. Teachers and, to a lesser extent, some school administrators, were not invested in using TCP as a resource to drive development because no one had asked them to use it in this way. We had created half of a new organizational capability – a better

way to measure performance – but without ownership from the "front lines," namely

teachers and all school administrators, we had not built the other half of the capability.

My misstep was both in not articulating the vision and in not encouraging others to take

responsibility for owning it. Although I lack the charisma of Carlzon and lacked the

positional authority and credibility at DSST to convince others across the organization to

act differently, I can surely do more in future roles to clearly articulate a way forward and

to push responsibility to others.

   Resist defensive routines – self-author and sort towards coaching. As noted

previously, my own defensive routines hampered the organization's ability to build a new

capability. Like so many others, I have a tendency to want to appear "as rational as

possible, by which we mean laying out clear-cut goals and then evaluating our own

behavior on the basis of whether or not we've achieved them" (Argyris, 1994, p. 80).

Faced with feedback about the way that teachers were experiencing TCP, I sorted for

tasks that could be accomplished, goals that could be set and achieved. This meant

choosing to act on technical changes – improvements to survey content, the instructional

rubric, and the compensation model – rather than engaging teachers in deep self-

reflection, a process that would have been more messy, with results that were less easily

quantified.

   Instead of sorting for technical changes, I need to "sort toward coaching," hearing

feedback "as potentially valuable advice from a fresh perspective rather than as an

indictment of how [I've] done things in the past" (Heen & Stone, 2014, p. 110). Acting

quickly on feedback is important for any leader; it demonstrates to others that you have

heard their feedback and you are responsive. However, after the quick wins occur, the

hard work begins. For me, acting on technical changes to the measurement function of TCP was only the first step of improvement. The next step, which is beyond the scope of the strategic project and this capstone, is to engage in the hard work of constructing the vision and empowering others to own TCP as a resource for development. I recognized the need for this next step only in the process of writing this capstone – sorting toward coaching by seeing teachers' perspectives on TCP not as an indictment of my failure to position the system as a developmental resource, but rather as an indication of the work left to do to realize TCP's developmental intent.

2. Implications for Site

Realizing TCP's developmental intent should remain a key priority for DSST Public Schools. Three years in the making when I joined DSST, the Teacher Career Pathway represented a significant effort by a growing system of schools to build a new organizational capability – specifically, the ability to "identify, develop, and retain effective teachers" via an evaluation system. This capstone can provide insight into the work ahead. Below, I offer three recommendations for DSST.

Decouple measurement and development systems. Educational researcher James Popham, generally a proponent of teacher evaluation, recently wrote, "Individuals who truly believe that a combined formative and summative teacher-evaluation effort can succeed are most likely to have recently arrived from outer space…. The only way to avoid such contamination is to keep the two enterprises separate" (2013, p. 18). Although hyperbolic, Popham raises an important point. Belief in coupling measurement and development is pervasive in K-12 education, including at DSST, owing not to interstellar travel, but to the expectations and directives raised by federal and state

legislation between 2009 and 2014, a point to which I will return later. The results of my strategic project suggest that teachers will primarily orient around evaluation systems as measurement tools if efforts are not also taken to construct aligned systems of coaching and support.

In order for DSST to create a systemic developmental resource, it must decouple measurement and development, building a professional learning system that parallels the robust measurement function of TCP. Several systems of schools have already started down this path. For example, District of Columbia Public Schools (DCPS), which launched its teacher evaluation system, IMPACT, in 2009, subsequently created a second system, LIFT, in 2012. As discussed in the RKA, IMPACT is explicitly a performance measurement system, producing one of five "effectiveness" ratings for teachers in DCPS. Teachers who receive ratings of "ineffective" and "minimally effective" are subject to separation from the district. "Effective" and "highly effective" teachers are recognized and rewarded. The purpose is clear: IMPACT measures performance. LIFT, on the other hand, is a developmental continuum. Based on multiple years of IMPACT effectiveness ratings, teachers move along a LIFT career pathway, with different opportunities for professional development and different observation processes at each stage. LIFT is what Hill and Grossman would call "a complementary system for instructional improvement" that does not "assume that evaluation systems built for accountability can serve dual purposes" (2013, p. 372).

Although TCP began as a system trying to "serve dual purposes," it does not have to remain as such. Home Office staff and school administrators could acknowledge that TCP's purpose is to measure performance and then, in a spirit of collaboration and

learning, work together with teachers to build a "complementary system" of coaching, professional development, and differentiated observation that helped to accelerate teachers' improved performance. Data produced by the TCP measurement system would still represent a point-in-time, summative evaluation of a teacher's progress on a developmental pathway and could be used in annual or semi-annual goal-setting conversations; however, formative data from regular coaching conversations and analysis of student interim assessments would guide movement on the pathway.

Building complementary systems would require a couple of organizational changes. First, DSST would have to decouple measurement and development roles at the school level. For reasons discussed earlier, observers cannot serve as both evaluator and coach if we expect observation processes to improve instruction. At YES Prep Public Schools, a charter management organization in Houston, Texas, that shares many characteristics with DSST, all first-year teachers have both an assigned coach and an evaluator. They are never the same person. DSST could extend this model to all teachers, a move with financial and staffing implications, to be sure, but one that may be required in order to effectively decouple measurement and development.

Second, in order to build complementary systems, DSST would have to form a Home Office team, even a cross-functional team, devoted to creating and maintaining two parallel systems. At present, responsibility for coaching and professional development is split across school leadership teams, teacher-leaders, a Home Office Academic team, and a Home Office Schools team. Constructing a shared vision and design for a complementary professional learning system may require an even greater level of investment and intentional collaboration among Home Office staff, school

leaders, and teachers, than what produced the original design for the Teacher Career Pathway. As I was writing this capstone, planning for such a team was underway, driven partly by my analysis of the strategic project. I am hopeful that these plans will come to fruition.

Foster the mental complexity of younger teachers. The results of my strategic project suggest a division between younger, less experienced teachers and older, more veteran teachers – with younger teachers more likely to orient vigorously around TCP measures as an indication of self-worth. I hypothesized that these subtle differences in perspective may be explained by theories of adult development. Although the sample sizes are very small and the results reflect only a year of perspectives on TCP, younger teachers seemed to be more "subject" to the system than were older, more veteran teachers. In order to decouple measurement and development, DSST may need to focus on building the mental complexity of younger teachers to understand measurement and development as distinct, but interrelated.

An example may help illustrate this implication. As described in the Introduction, TCP does not attempt to label teachers as "effective" and "ineffective," judging ineffective teachers as unfit for employment, as many evaluation systems across the country do. DSST's status as a charter school organization gave it the flexibility to define evaluation categories differently than what had been imagined under SB 10-191. This is a special feature of TCP. "Novice" does not mean "ineffective"; instead, it means that one is at the beginning of a learning curve. However, my experience discussing TCP with scores of DSST teachers, particularly younger teachers, indicates that many do not view placements in this way. At one feedback session in November 2014, a teacher in

his mid-20's said, "If I placed at Novice, I would quit because the sign would be so clear that I needed to get out of teaching." When I asked him how he came to this conclusion, he explained, "Novice means you're bad; you just are." Probing further, I reframed TCP placements for this teacher, asking him to reflect on the difference between a label of "ineffective" and one of "Novice." This is more than semantics, I suggested. "That's interesting," the teacher said. "I never thought about it in that way."

Helping others to develop mental complexity is nonlinear work. It does not produce clearly defined outcomes, a fact that can be challenging for achievement-oriented organizations. However, if DSST is to continue to follow a "rare human capital model" of employing young, well-educated teachers who are themselves highly achievement-oriented, it must find ways to foster mental complexity among younger teachers, helping them to understand evaluation measures not as fixed labels, but as fair reflections of a stage of professional growth (Wilson, 2008, p. 2).

Nesting a robust coaching model within the complementary development system described under the first implication could provide an infrastructure for building mental complexity. Coaches could be trained to support teachers' growth through structured inquiry and other protocols intended to elicit reflections on practice and to surface the deeply held assumptions and beliefs whose interrogation promotes enduring changes in behavior. Organizational defensive routines may resist and dismiss this type of coaching as a distraction from the core work of coaching teachers to design better lesson plans or effectively execute a gradual release of responsibility. However, skilled coaches do both: they train teachers on core instructional practices and build in others the mental dexterity to take as *object* that which one previously took as *subject*. Coaching would look

differently depending on the age and developmental stage of each teacher, but all coaching would include a concurrent focus on technical and adaptive skill building.

Engage experienced teachers in collaborative team learning alongside Home Office staff and school administrators. The RKA began with a discussion of Achievement First (AF), a network of charter schools in the Northeast U.S. In Rosenberg's view (2012), AF experiences a vicious cycle where experienced teachers and leaders who have the capacity to build systems that support deep teacher development leave the organization, feeling stagnated and frustrated because such systems do not exist. DSST faces a similar challenge. Teacher retention plummeted below 70% in 2013-2014. In 2014-2015, four of nine DSST School Directors were new to their jobs. At the school where the teacher vowed that he would quit with a placement of Novice, four of five members of the school leadership team were new to their positions.

In order to construct a complementary development system and coaching model that supports continuous teacher growth, the organization needs teachers and leaders with experience and expertise in the DSST instructional model. It needs teachers and leaders who will remain in their jobs. This is especially true of instructional leaders and the Lead and Master teachers who have demonstrated the greatest levels of expertise. In my residency work outside of TCP, I helped to lead an organization-wide strategic initiative intended to increase career sustainability through changes to team norms and culture, improvements to schedule, staffing, and workload, and building of leadership capacity. Given the large number of teachers and leaders in their first few years with the organization, this initiative was increasingly focused on the needs of "new-to-DSST"

employees. Equally important, however, was attention to veteran teachers and leaders in constructing a development system.

Going forward, DSST might consider creating hybrid roles for Master, and potentially Lead, teachers, allowing them to serve as formal coaches to teachers at earlier stages on the pathway and collaborating with school administrators and Home Office team members to build the complementary development system. Such hybrid roles might include part-time teaching and part-time leadership responsibilities, with costs offset by expansion of the network's Apprentice Teacher program or possible reduction of school-level administration by one full-time equivalent position. These roles might prevent stagnation and burnout and leverage the experience and expertise of Master and Lead teachers.

This collaborative work with more experienced teacher-leaders could draw insight from learning processes on the core TCP team. Among that team, I created an environment where members were able to learn from the team's work-in-progress and make modifications based on that learning. The team operated flexibly, refining the design as implementation unfolded. However, the TCP team was composed entirely of system designers – the Home Office managers and directors responsible for the functional work of implementation. School leaders became part of the learning process, but teachers, despite being on the original TCP Working Group, did not participate fully in the effort to evolve the new system. As DSST looks to construct a complementary development system aligned to TCP, the organization might ask not only how experienced Lead and Master teachers could contribute through feedback and occasional

creation of work products (e.g. coaching protocols and materials), but also how these teachers could join the cross-functional development and implementation team.

For example, by prioritizing hybrid roles for Lead and Master teachers, DSST could allow these teachers to actively and completely participate in a "professional learning design and implementation" team, with responsibility for creating and executing TCP-aligned coaching and support, alongside Home Office managers and directors and, potentially, school administrators. Working together in a climate of psychological safety, acknowledging individual and collective errors, and learning through inquiry, this team might serve as a catalyst for greater trust and understanding between teachers and administrators and help to flatten DSST's hierarchy, while simultaneously producing an aligned development model that reflects the best thinking and collective investment of key stakeholders.

Creating space for this type of team learning among teachers and administrators will be unusual at DSST, as it would be within most school systems. However, the identification of Lead and Master teachers under the TCP system and the organization's initial movement towards creating an aligned system of support provide a unique opportunity for DSST to be a leader in opening space for what Rick Hess calls "cage-busting teachers," those who "champion excellence, identify important problems, offer concrete solutions, and bring those solutions to life" in roles that put teachers at the table with school and system leaders (Hess, 2015). The mode of team learning discussed in this capstone could offer a novel template for learning among teachers and administrators. Moreover, it could enliven and challenge experienced teachers in a way that further prevents burnout and stagnation.

3. Implications for Sector

Educational policy is notoriously ineffective at achieving its desired outcomes, American history littered with policy reforms that failed to alter practice (Elmore, 1996; Cohen & Moffitt, 2009). This capstone set out to explore the implementation of an evaluation system motivated by one of the nation's highest-profile pieces of recent educational policy, Colorado's Senate Bill 10-191. DSST's experience during the first year of implementation has important implications for policymakers in not only Colorado, but also nationwide.

Align the instruments of policy with the instruments of practice. Where policy "works," one finds substantial "coherence among the instruments of policy at the level of practice," according to Cohen and Hill (2001), who undertook one of the most exhaustive studies ever to examine the impact of a state education policy on school-level practice (p. 7). "Standards, assessments, and accountability are more likely to succeed," Cohen and Hill write, "if they are accompanied by extended opportunities for professional learning that are grounded in practice" (2001, p. 10).

The results of the strategic project lend support to Cohen and Hill's assertion. The instruments undergirding the measurement purpose of SB 10-191 – including mandates for annual evaluation, use of student achievement data to measure performance, and criteria for assigning educators to evaluation categories – were well-aligned with the instruments of practice already in use at DSST. Indeed, although largely subjective, annual evaluations already occurred. Teachers and leaders generally believed that student achievement results were a fair reflection of a teacher's effectiveness and they assiduously collected and analyzed such results. A common teaching rubric was already

used for classroom observations and survey data were in use as a formative and summative resource across the network. In short, DSST possessed the instruments of practice to deliver on the legislation's measurement intent. The challenge was to bring these instruments together into a single system, a new organizational capability. As discussed in the Analysis section, DSST stakeholders generally felt invested in developing this capability, even if they had fundamentally different stakes in its development.

For school systems that lacked the instruments of practice, the legislation tried to produce coherence by funding the creation of "assessments, processes, tools, and policies that [could] be used to develop evaluation systems" (Engdahl, 2010). The Colorado Department of Education created a model evaluation system that could be adopted by any district and, through a non-profit partner, offered training on development and implementation of evaluation systems. Nevertheless, the gap was profound between the instruments of policy and those of practice found in most Colorado school systems. In Spring 2014, districts successfully petitioned Colorado lawmakers to delay enforcement of the student achievement provision of SB 10-191, saying that they were unable to meet the timeline for implementation (Engdahl, 2014). DSST has been one of the few Colorado school systems to fully implement the legislation according to its original schedule. Almost everywhere, the aspiration of the legislation to measure the performance of Colorado's educators outstripped the ability of the sector to deliver.

Perhaps more significantly, when it came to the development purpose of SB 10-191, the instruments of policy were misaligned with the instruments of practice. The legislature told Colorado's school systems to tie evaluation to professional learning

opportunities, but provided no guidance, no tools, and, importantly, no funding that would allow schools to link evaluation with professional learning. Instead, design of evaluation systems occurred in the midst of five straight years of state-level budget cuts (Simpson & Torres, 2014). The creation of a complementary development system and robust coaching model currently being contemplated by DSST would be funded largely through private donations. In essence, the development purpose of SB 10-191 has been an unfunded mandate – no instruments of policy to support the level of practice. If policymakers expect school systems to align measurement and development, they must provide financial resources, capacity-building guidance and toolkits to create this alignment.

Decouple measurement and development at the level of policy. This capstone raises the question about whether teacher evaluation policy, in Colorado and beyond, simply confuses the sector by conflating measurement and development. The origins of this fraught coupling lie with the federal Department of Education, which, through 2009's Race to the Top grant competition and the ongoing program to grant states waivers from provisions of the Elementary and Secondary Education Act (ESEA), incentivized states to create teacher evaluation policy intended both to measure and develop teachers. A review of the guidelines for ESEA flexibility illustrates this tight coupling. The Department has issued six criteria for states' teacher evaluation policies. Four of these criteria have a measurement purpose; two have a development purpose (Appendix G).

The results of my strategic project suggest that schools may struggle to create evaluation systems that both measure and develop teachers. As discussed earlier, a better approach may be to create complementary systems. Policymakers can play a role by

incentivizing the construction of complementary systems, removing the development purpose from current teacher evaluation statute and passing new legislation whose sole aim is to motivate schools to align professional learning with new and emerging evaluation systems. For instance, states might create grant competitions that provide funding to school systems in order to design professional learning opportunities, including coaching and traditional professional development, to teachers; evaluation of the program would occur through rigorous research that sought to identify the impact of these opportunities on teachers' growth over time, as measured by evaluations.

At the federal level, Congress might consider legislation modeled partly on the Teacher Incentive Fund (TIF), a discretionary grant program that provides funding to school systems for design and implementation of performance-based compensation. Recent versions of the TIF program have provided guidance to grantees on the alignment between evaluation measures and performance bonuses, linking newly created evaluation systems with federally funded systems to differentiate teacher compensation (Max et al. 2014). Congress might consider similar legislation that incentivizes the creation of professional learning systems aligned to evaluation systems, thus building capacity for both measurement and development.

Heed the "competitive imperative of learning." For over twenty years, researchers have called for schools to become "learning organizations," defined by shared vision, team dialogue and inquiry, collaboration, and systems thinking (Senge, 2000; Fullan, 1995). This capstone underscores both the importance and the challenge of building learning cultures within systems of schools. My experience leading the cross-functional TCP team suggests that it is possible to create a culture of learning within a

school central office and that this culture may even extend out to include school administrators. However, the structure of schooling, based on a "prevailing system of management," complicates efforts to cultivate a culture of learning across an entire school system – whether the origin of that effort is a single classroom or a central office (Senge, 2006, p. xi). Nevertheless, this capstone suggests some lessons for cultivating a learning culture across a system of schools.

Take time for dialogue with a purpose. In her article, "The Competitive Imperative of Learning," Edmondson writes that "the only way to achieve and sustain excellence is for leaders to insist that their organization invest in the slack time and resources that support [disciplined reflection]" (2008, p. 67). Between July and October 2014, I devoted many hours of the TCP team's meeting time to dialogue that built shared ownership, fostered collaboration, and promoted cross-functional understanding. School Directors also invested significant time in dialogue during monthly network-wide Leadership Team meetings and in regular conversations with either the Director of Human Capital or me. These conversations surfaced divergent viewpoints and assumptions and opened space for inquiry and learning, yet remained rooted in a common purpose – how might we build a career pathway system that identifies, develops, and retains great teachers? System-level leaders seeking to build a learning culture within their organizations must invest in time for dialogue.

Collaborate with teachers. The effort to invest stakeholders in the implementation of TCP as a learning endeavor faltered with DSST's teachers. While I have argued that the high, personal stakes of evaluation for teachers complicated this effort, so too did my own use of communication tools that blocked learning. Cultivating a learning culture

across schools means complementing staff surveys and town hall sessions with extended opportunities for teachers and administrators to collaborate together in a climate of psychological safety and inquiry, not around the bargaining table, but in dialogue about common challenges and shared opportunities.

Model a learning stance. Creating a holding environment defined by organizing-to-learn principles required me to model the types of behaviors that I hoped to instill in others. This is an important lesson for system-level leaders. Given a bureaucratic structure based implicitly on command-and-control, leaders must *go out of their way* to demonstrate to others a willingness to acknowledge and learn from failure, build shared ownership, value diverse viewpoints, and celebrate collaboration. Taking a learning stance is often unnatural for leaders who "embrace the language" of learning, but "fail to see how firmly mired" in prevailing management practices their words and actions really are (Argyris, 1994, p. 83). System-level leaders must take time for personal reflection on their own habits and defensive routines. When leaders take time to do this work and genuinely commit to a learning stance, they begin to dismantle the barriers that separate central office from schools, school leaders from teachers, and teachers from students. This is difficult work. It is indispensable work.

Conclusion

On a warm evening in January 2015, Colorado State Senator Michael Johnston, the chief architect of Senate Bill 10-191, stood before a room of DSST teachers, their family members, school leaders, and other supporters, and said that the evening represented "the culmination of a dream." The occasion was DSST's first-ever "Night to

Celebrate Exemplary Teaching," a dinner honoring teachers who had reached the Lead

and Master stages on the Teacher Career Pathway in 2014, as well as recipients of special

honors aligned to the pathway. Over the course of the evening's program, which I

designed, DSST students offered personal tributes to each of the nine Master teachers

and, through video, colleagues and students took the audience on a journey to understand

the impact that these individual teachers had on students' lives.

Kicking off the evening, Senator Johnston said:

> I always dreamed this night would happen, which is when we started the idea of
> teacher evaluation reform, this was the biggest hope. The hope was that if you
> built the system that gave people really clear feedback about their performance
> and really clear targets about where to improve, what you would do would be to
> unleash the talent of incredibly talented educators and help us figure out what
> makes them so incredible.
>
> We would create a system where you all led the profession as opposed to
> policymakers or district leaders or the rest trying to lead. It would allow us to get
> out of the way and allow great teaching to reign. I think people got confused that
> the conversation was about poor performance when the conversation was really
> about unleashing outstanding performance and we were just smart enough to
> know that we didn't have all the answers.
>
> So, this really is the culmination of a dream. The best and highest intent of the
> legislation that we passed – and now 24 other states have tried to replicate – was
> that we might one night gather in rooms like this where our strongest and most
> compelling teachers led the way.

Johnston was right on several counts. The measurement function of teacher

evaluation has the potential to identify "our strongest and most compelling teachers,"

systematically recognizing them for their contributions to children's lives in a way that

we, as a country, have never recognized teachers before. Indeed, following that January

evening, many honored teachers said that it was the greatest public appreciation that they

had experienced in their careers. Such recognition, symbolic and financial, has the

potential to retain great teachers in the profession and at DSST. As one teacher wrote on the morning after the event, "I left feeling very appreciated, honored, loved, and fortunate. I've had a wonderful life with DSST these past seven years, and look forward to many more in the future." Moreover, by giving teachers feedback on their performance and "clear targets about where to improve," teacher evaluation provides critical data that can be used to help teachers improve their practice.

However, there is a great distance between the availability of performance data and the use of such data to grow and develop. Across the nation, "people got confused" about evaluation because educational leaders – from the local to the national level – failed to articulate a vision for how evaluation was going to develop teachers, for how it was going to unleash outstanding performance. This is the work ahead not only for DSST, but also for the sector.

Reviewing Achievement First's Teacher Career Pathway in 2011, Rachel Curtis wrote that leaders must "commit first and foremost to learning from their efforts and continually refining their design and implementation based on the learning" (p. 22). This capstone underscores the importance of approaching teacher evaluation reform as a learning endeavor, requiring flexibility, psychological safety, and adaptation. However, learning by leaders alone will not position evaluation as a resource for teacher growth and development. As the experience at DSST highlights, leaders may learn, but without a collaborative process that includes teachers in the dynamic, evolving work and builds complementary systems for development, evaluation is unlikely to achieve its lofty developmental goals. System designers must commit first and foremost to learning alongside system users. The sector is indeed at the beginning of the learning curve.

Bibliography

Achievement First. (2011). *AF Teacher Career Pathway Executive Summary*. Retrieved 4 September 2014 from http://tntp.org/assets/tools/AF%20Teacher%20Career%20Pathway%20Executive%20Summary%20Both%20Regions%20TSLT_0311.pdf.

Angrist, J. D., Pathak, P. A. & Walters, C. R. (2011). *Explaining charter school effectiveness*. NBER Working Paper No. 17332. National Bureau of Economic Research.

Argyris, C. (1971). *Management and organizational development; the path from XA to YB*. New York: McGraw-Hill.

Argyris, C. (1977). Double loop learning in organizations. *Harvard Business Review*, 55(5), 115-125.

Argyris, C. (1991). Teaching smart people how to learn. *Harvard Business Review*, 69(3), 99-109.

Argyris, C. (1994). Good communication that blocks learning. *Harvard Business Review,* 72(4), 77-85.

Argyris, C. & Schön, D. (1978). *Organizational learning: a theory of action perspective*. Reading, MA: Addison-Wesley.

Barth, R. (2006). Improving relationships within the schoolhouse. *Educational Leadership*, 63(6), 8-13.

Cameron, K.S. & Quinn, R.E. (1999). *Diagnosing and changing organizational culture*. Reading, MA: Addison-Wesley.

Carlzon, J. (1995). Jan Carlzon. *Harvard Business School*. Case 8-489-046.

Carroll, V. (2014). It's time to give DSST schools their due. *The Denver Post*. 2 September 2014. Retrieved 28 December 2014 from http://www.denverpost.com/carroll/ci_26454444/its-time-give-dsst-schools-their-due.

Chubb, J.E. & Moe, T.M. (1990). *Politics, markets, and America's schools*. Washington, D.C.: The Brookings Institution.

Chute, E. (2014). Pittsburgh teachers fare well in new rating system. *Pittsburgh Post-Gazette*. 12 June 2014. Retrieved 31 August 2014 from http://www.post-gazette.com/news/education/2014/06/13/Pittsburgh-teachers-fare-well-in-new-rating-system/stories/201406120300.

Christensen, C. & Kaufman, S. (2008). *Module Note – Assessing Your Organization's Capabilities: Resources, Processes, and Priorities*. 607-014.

Clark, K.B., & S. C. Wheelwright (1992). "Organizing and Leading 'Heavyweight' Development Teams." *California Management Review*, 34(3), 9–28.

Clark, K.B. & Fujimoto, T. (1991). *Product development performance: strategy, organization, and management in the world auto industry.* Cambridge, MA: Harvard Business Review Press.

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

Cohen, D.K. & Moffitt, S.L. (2009). *The ordeal of equality: did federal regulation fix the schools?* Cambridge, MA: Harvard University Press.

Colorado Department of Education. (2014). Memo. 13 June 2014. *Updated Guidance: How S.B. 10-191 (The Great Teachers and Leaders Bill), applies to charter schools and innovation schools?* Retrieved 4 December 2014 from http://www.cde.state.co.us/cdechart/updatedguidanceonsb-191.

Corcoran, S.P. (2010). *Can teachers be evaluated by their students' test scores? Should they be? The use of value-added measures in teacher effectiveness in policy and practice*. Education Policy Action Series. Providence, RI: Annenberg Institute for School Reform.

CREDO. (2013). *National charter school study 2013*. Stanford, CA: Center for Research on Education Outcomes.

Curtis, R. (2011). *Achievement First: developing a teacher performance management system that recognizes excellence*. Washington, D.C.: The Aspen Institute.

Dee, T. & Wyckoff, J. (2013). *Incentives, selection, and teacher performance: evidence from IMPACT*. NBER Working Paper No. 19529. National Bureau of Economic Research.

Di Carlo, M. (2010). The 5-10 percent solution. *Shanker Blog*. 16 December 2010. Retrieved 28 December 2014 from http://shankerblog.org/?p=1473.

Di Carlo, M. (2014). Revisiting The Widget Effect. *Shanker Blog*. 4 March 2014. Retrieved 31 August 2014 from http://shankerblog.org/?p=9598.

Dobbie, W. & Fryer, R. G., Jr. (2011). *Getting beneath the veil of effective schools: Evidence from New York City.* NBER Working Paper No. 17632. National Bureau of Economic Research.

Doherty, K.M. & Jacobs, S. (2013). "Connect the Dots: Using Evaluations of Teacher Effectiveness to Improve Policy and Practice." *National Council on Teacher Quality*. Retrieved 22 November 2014 from http://www.nctq.org/dmsView/State_of_the_States_2013_Using_Teacher_Evaluations_NCTQ_Report.

Doyle, D. & Han, J.G. (2012). *Measuring teacher effectiveness: a look "under the hood" of teacher evaluation in 10 sites*. Retrieved 4 September 2014 from http://50can.org/sites/50can.org/files/Measuring%20Teacher%20Effectiveness.pdf.

DSST Public Schools. (2014). *TCP Guidebook*. Provided by DSST Public Schools.

Duncan, A. (2009). *Partner in reform: remarks of Arne Duncan to the National Education Association.* 2 July 2009. Retrieved 31 August 2014 from http://www2.ed.gov/news/speeches/2009/07/07022009.html.

Edmondson, A. (2008). The competitive imperative of learning. *Harvard Business Review,* 86(7), 60–67.

Edmondson, A. (2012). *Teaming: how organizations learn, innovate, and compete in the knowledge economy*. San Francisco, CA: Jossey-Bass.

Elmore, R. F. (1996). Getting to scale with good educational practice. *Harvard Educational Review*, 66 (1), 1–27.

Engdahl, T. (2010). Inside Senate Bill 10-191. *Chalkbeat Colorado.* 17 May 2010. Retrieved 21 February 2015 from http://co.chalkbeat.org/2010/05/17/inside-senate-bill-10-191/.

Engdahl, T. (2014). Bill proposes delay in using student growth for evaluations. *Chalkbeat Colorado*. 6 March 2014. Retrieved 21 February 2015 from http://co.chalkbeat.org/2014/03/06/bill-proposes-delay-in-using-student-growth-for-evaluations/.

Fullan, M. (1995). The school as learning organization: Distant dreams. *Theory into Practice*, 34(4), 230–235.

Gardner, H. K. (2012). Coming through when it matters most. *Harvard Business Review,* 90(4), 82-91.

Green, E. (2014). *Building a better teacher: how teaching works (and how to teach it to everyone).* New York: W.W. Norton & Co.

Hanushek, E. A. (2011). Valuing teachers: how much is a good teacher worth? *Education Next,* 11(3), 40-45.

Harris, E.A. (2014). Critics question high ratings on New York State teacher evaluations amid poor test scores. *The New York Times*. 28 August 2014. Retrieved 31 August 2014 from http://www.nytimes.com/2014/08/29/nyregion/new-york-state-releases-localized-teacher-evaluations.html.

Heen, S. & Stone, D. (2014). Find the coaching in criticism. *Harvard Business Review*, 92(1), 108-111.

Heifetz, R. A. (1994). *Leadership without easy answers*. Cambridge, MA: Belknap Press of Harvard University Press.

Hess, F. M. & Higgins, M. (2009). *The challenges for charter schools: replicating success: education outlook no. 4*. American Enterprise Institute for Public Policy Research.

Hess, F.M. (2015). A breakout role for teachers. *Education Next*. Retrieved 28 March 2015 from http://educationnext.org/breakout-role-teachers/.

Hill, H. C. & Grossman, P. (2013). Learning from teacher observations: challenges and opportunities posed by new teacher evaluation systems. *Harvard Educational Review* 83(2): 371-384.

Hoxby, C. M. & Murarka, S. (2009). *Charter schools in New York City: Who enrolls and how they affect their students' achievement.* NBER Working Paper No. 14852. National Bureau of Economic Research.

Institute for Education Sciences. (2014). *State requirements for teacher evaluation policies promoted by Race to the Top.* NCEE Evaluation Brief. Retrieved 31 August 2014 from http://ies.ed.gov/ncee/pubs/20144016/pdf/20144016.pdf.

Johnston, M. & Dallman, K. (2014). "Giving Colorado's teacher evaluation bill time to succeed." The Denver Post. 6 March 2014. Retrieved 22 November 2014 from http://www.denverpost.com/opinion/ci_25288155/giving-colorados-teacher-evaluation-bill-time-succeed?source=infinite.

Kegan, R. & Lahey, L. L. (2009). *Immunity to change: How to overcome it and unlock potential in yourself and your organization*. Boston, MA: Harvard Business Press.

Kegan, R., Lahey, L.L., Fleming, A. & Miller, M. (2014). Making business personal. *Harvard Business Review*, 92(4), 44-52.

Kuwashima, K. (2013). Three Footnotes to "Heavyweight Product Manager." *Annals of Business Administrative Science*, 12(2013), 265-276.

Marzano, R. J. (2012). The two purposes of teacher evaluation. *Educational Leadership*, 70(3),14-19.

Max, J., Constantine, J., Wellington, A., Hallgren, K., Glazerman, S., Chiang, H., Speroni, C. (2014). *Evaluation of the Teacher Incentive Fund: implementation and early impacts of pay-for-performance after one year*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

McGregor, D. (1960). *The human side of enterprise*. New York: McGraw-Hill.

McTighe, J. & Wiggins, G. P. (1999). *Understanding by design handbook*. Alexandria, VA: Association for Supervision and Curriculum Development.

Mehta, J. (2013). *The allure of order: High hopes, dashed expectations, and the troubled quest to remake American schooling*. New York: Oxford University Press.

Merseth, K. K. (2009). *Inside urban charter schools: Promising practices and strategies in five high-performing schools*. Cambridge, MA: Harvard Education Press.

Meyerson, D. (2001). "Radical Change, The Quiet Way," *Harvard Business Review*, 79(9), 92-100.

National Association of Charter School Authorizers & Charter School Growth Fund. (2014). *Replicating quality: policy recommendations to support the replication and growth of high-performing charter schools and networks*. Retrieved 13 September 2014 from http://nacsa.mycrowdwisdom.com/diweb/catalog/item/id/194990/q/n=9&c=82.

NewSchools Venture Fund. (2010). *A highly effective teacher in every classroom: creating talent development systems that drive instructional excellence*. Retrieved 4 September 2014 from http://www.newschools.org/files/talent-development-tool.pdf.

Ouchi, W. G. (1981). *Theory Z: How American business can meet the Japanese challenge*. Reading, MA: Addison-Wesley.

Owens, R. G. & Valesky, T. C. (2011). *Organizational behavior in education: Leadership and school reform*. Boston: Pearson.

Papay, J. (2012). Refocusing the debate: assessing the purposes and tools of teacher evaluation. *Harvard Educational Review,* 82(1),123-141.

Popham, W.J. (2013). *Evaluating America's teachers: mission possible?* Thousand Oaks, CA: Corwin.

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). *Organizing for quality in education: individualistic and systemic approaches to teacher quality*. Unpublished doctoral dissertation. Ann Arbor, MI: Michigan University.

Rotherham, A. & LiBetti Mitchel, A. (2014). *Genuine progress, greater challenges: A decade of teacher effectiveness reforms*. Washington, D.C.: Bellwether Education Partners.

Schein, E.H. (1993). On dialogue, culture, and organizational learning. *Organizational Dynamics,* 22(2), 40-51.

Schein, E.H. (1996). Three cultures of management: the key to organizational learning. *Sloan Management Review,* 38(1), 9-20.

Senate Bill 10-191. (2010). General Assembly of the State of Colorado.

Senge, P. M. (2000). *Schools that learn: A fifth discipline fieldbook for educators, parents, and everyone who cares about education*. New York: Doubleday.

Senge, P.M. (2006). *The fifth discipline: the art and practice of the learning organization, revised edition.* New York: Doubleday.

Simpson, K. & Torres, Z. (2014). Colorado teacher evaluations still face major hurdles after first year. *The Denver Post*. 28 June 2014. Retrieved 21 February 2015 from http://www.denverpost.com/news/ci_26054448/colorado-teacher-evaluations-still-face-major-hurdles-after.

Taylor, E. S., & Tyler, J. H. (2011). *The effect of evaluation on performance: Evidence from longitudinal student achievement data of mid-career teachers* (NBER Working Paper No. 16877). Washington DC: National Bureau of Economic Research.

Taylor, E.S., & Tyler, J.H. (2012). Can teacher evaluation improve teaching? *Education Next*, 12(4). Retrieved 10 February 2015 from http://educationnext.org/can-teacher-evaluation-improve-teaching/.

TNTP. (2010). *Teacher evaluation 2.0*. TNTP.

Toch, T. (2009). Charter-management organizations: Expansion, survival, and impact. *Education Week,* 29(9), 27-32.

Tucker, A.L., Edmondson, A.C., & Spear, S. (2002). When problem solving prevents organizational learning. *Journal of Organizational Change Management*, 15(2), 122-137.

Watkins, M. (2009). New leader? Get early wins. *Harvard Business Review*. Retrieved 8 February 2015 from https://hbr.org/2009/01/new-leader-get-early-wins/.

Weber, M. (1978). *Economy and Society.* Berkeley, CA: University of California Press.

Weinberg, D. et al. (2009). *The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness.* TNTP.

Williams, D. (2010). *Real leadership: helping people and organizations face their toughest challenges, ReadHowYouWant edition.* Sydney, Australia: Accessible Publishing Systems.

Wilson, S. F. (2009). *Success at scale in charter schooling. education outlook no. 3*. American Enterprise Institute for Public Policy Research.

Appendix A: TCP Team charter, July 2014

<u>Teacher Career Pathway</u>

At DSST, identifying, developing and retaining effective teachers are the most important efforts we can take to increase student achievement and realize our mission and our team's career satisfaction. The Teacher Career Pathways (TCP) is our organization's strategy, developed by a collaborative group of teachers, leaders, and staff, to realize this critical priority. TCP provides differentiated professional development, responsibilities, recognition, and compensation for all teachers in hopes of sustaining long-term student achievement and career satisfaction.
TCP evolves and improves over time based on feedback from teachers and leaders.

As a TCP team, we strive to:
- Continue to build, modify, and improve TCP from FY14 to FY15 so that the entire DSST team has a resource and tool to support ongoing teacher development across the network
- Communicate clearly and consistently about design and key decisions, speaking with one voice about all things TCP
- Work collaboratively to understand individual responsibilities as well as their fit within the larger system
- Continue to leverage internal team member strengths and foundational work of the development and pilot phases in order to accomplish the above.

<u>Team Structure</u>

We commit to open and transparent communication, where key decisions are made through the combined thinking and expertise of the team during regular meetings and where information is shared widely through virtual channels in order to facilitate broad understanding of the larger system.
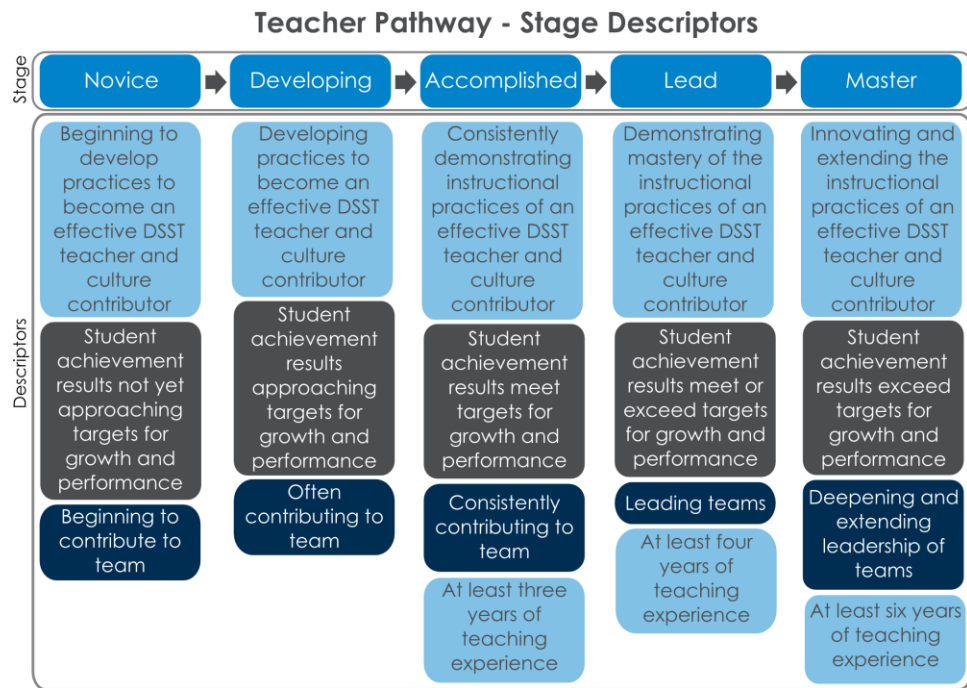
<u>Roles & Responsibilities</u>
- Academics
  - Continue to hone assessment strategy across all grades and subject areas (Director of Curriculum and Instruction)
  - Oversee ongoing revisions to the TE Rubric and ensure that formal observers are normed and able to officially review team members (Director of Curriculum and Instruction)
  - Oversee norming on survey implementation and development of feedback cultures (Director of Schools)
  - Design and execute professional development for all teachers (Senior Manager of Professional Development)
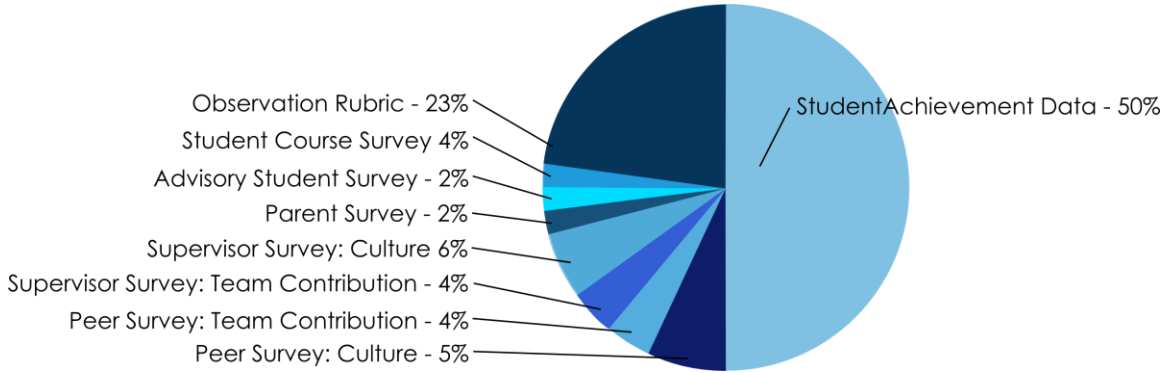- Data

- o Collect and analyze all data related to TCP, including student achievement, observational data, and peer, parent, and student surveys (Senior Manager of Data and Assessment)
  - o Present data in actionable form that drives decision-making (Senior Manager of Data and Assessment)
  - o Lead, design, and modify data analysis strategy (Senior Manager of Data and Assessment)
- Education technology
  - o Ensure that teachers and School Directors have access to clear, engaging reports and analysis that support ongoing feedback and development (Senior Manager of Education Technology)
  - o Ensure integrity and accuracy of reported data (Senior Manager of Education Technology)
  - o Oversee data warehousing and longitudinal organization to ensure scalability and reliability (Senior Manager of Education Technology)
- Operations
  - o Manage collection of all CTI data and optimize functionality of data collection tools (Director of Operations)
  - o Provide regular updates regarding completion of surveys and performance reviews (Director of Operations)
- Finance
  - o Provide regular updates regarding long-term financial implications of TCP design and implementation, as well as budgetary information to guide short-term decision making (Director Finance)
  - o Process payroll to reflect compensation determined through TCP (Director of Finance)
- Human capital
  - o Develop clear and consistent communication strategy, including via design of training and other communications (Director of Human Capital/Resident)
  - o Ensure that all components of the TCP system – operations, finance, academics, data, and education technology – operate seamlessly (Resident)
  - o Develop and maintain TCP budget (Resident)
  - o Create and facilitate long-term engagement strategy that results in improvements based on feedback from teachers and other stakeholders (Resident)
  - o Manage systems to ensure differentiated responsibilities, recognition, compensation, and professional development for teachers (Director of Human Capital)
  - o Coordinate determination of compensation levels and provide such information to Finance for payroll processing (Payroll Coordinator)

Appendix B: Teacher Career Pathway Framework and Stage Descriptors

Teacher Career Pathway Framework

The Teacher Career Pathway Framework outlines the measures used to assess the effectiveness of teachers in DSST Public Schools. All teachers at DSST will be evaluated on student achievement results, culture, contributions to school and network teams, and instructional practice.

Student Achievement	Culture Contributions	Team (Leadership)	Instructional Practice
Achieving student growth as measured by **student achievement results**	Living and upholding DSST Core Values as evaluated in **feedback surveys**	Contributing to the team and the organization as measured by **feedback surveys**	Implementing DSST Core Instructional Practices as documented by the **TE Rubric and student survey**
Tools: ☒ IAs ☒ EPAS ☒ MAP Exams ☒ Internally or externally created assessments for non-tested subjects	**Tools:** ☒ Supervisor Evaluation ☒ Peer Surveys ☒ Parent Surveys ☒ Student Advisory Surveys	**Tools:** ☒ Supervisor Evaluations ☒ Peer Surveys	**Tools:** ☒ DSST Instructional Rubric ☒ Student Course Surveys

Teacher Pathway - Stage Descriptors

Stage	Novice	Developing	Accomplished	Lead	Master
Descriptors	Beginning to develop practices to become an effective DSST teacher and culture contributor	Developing practices to become an effective DSST teacher and culture contributor	Consistently demonstrating instructional practices of an effective DSST teacher and culture contributor	Demonstrating mastery of the instructional practices of an effective DSST teacher and culture contributor	Innovating and extending the instructional practices of an effective DSST teacher and culture contributor
	Student achievement results not yet approaching targets for growth and performance	Student achievement results approaching targets for growth and performance	Student achievement results meet targets for growth and performance	Student achievement results meet or exceed targets for growth and performance	Student achievement results exceed targets for growth and performance
	Beginning to contribute to team	Often contributing to team	Consistently contributing to team	Leading teams	Deepening and extending leadership of teams
			At least three years of teaching experience	At least four years of teaching experience	At least six years of teaching experience

Appendix C: Distribution of evaluative components

Observation Rubric - 23%
Student Course Survey 4%
Advisory Student Survey - 2%
Parent Survey - 2%
Supervisor Survey: Culture 6%
Supervisor Survey: Team Contribution - 4%
Peer Survey: Team Contribution - 4%
Peer Survey: Culture - 5%
Student Achievement Data - 50%

Appendix D: Teacher Career Pathway Placement Matrix

Teacher Career Pathway Matrix

Culture, Team, and Instruction (CTI)
- 5 categories (I,II,III,IV,V)
- Wider range in performance

Student Achievement Data
- 3 categories (A,B,C)
- Smaller differences in performance

Years of Experience
- Full-time teaching experience as a teacher of record

8

127

Appendix E: Placement conversation toolkit for School Directors

Appendix F: Distribution of teachers on TCP

Teachers at each stage on Teacher Career Pathway, October 2014		
Stage	**#**	**%**
Novice	7	7%
Developing	28	28%
Accomplished	38	38%
Lead	17	17%
Master	9	9%
TOTAL	99	100%

Appendix G: U.S. Department of Education. ESEA Flexibility, Updated June 7, 2012. Quoted in Popham, 2013, p. 5.

Annotation by author: **Measurement purposes in bold**. Development purposes underlined.

To receive this flexibility, an SEA and each LEA must commit to develop, adopt, pilot, and implement, with the involvement of teachers and principals, teacher and principal evaluation and support systems that: (1) will be used for continual improvement of instruction; **(2) meaningfully differentiate performance using at least three performance levels**; **(3) use multiple valid measures in determining performance levels, including as a significant factor data on student growth for all students (including English Learners and students with disabilities), and other measures of professional practice (which may be gathered through multiple formats and sources, such as observations based on rigorous teacher performance standards, teacher portfolios, and student and parent surveys); (4) evaluate teachers and principals on a regular basis;** (5) provide clear, timely, and useful feedback, including feedback that identifies needs and guides professional development; **and (6) will be used to inform personnel decisions**. An SEA must develop and adopt guidelines for these systems, and LEAs must develop and implement teacher and principal evaluation and support systems that are consistent with the SEA's guidelines. To ensure high-quality implementation, all teachers, principals, and evaluators should be trained on the evaluation system and their responsibilities in the evaluation system. As part of developing and implementing these evaluation and support systems, an SEA must also provide student growth data on current students and the students taught in the previous year to, at a minimum, teachers of reading/language arts and mathematics in grades in which the State administers assessments in those subjects in a manner that is timely and informs instructional programs.