INCENTIVES IN THE CLASSROOM: PERFORMANCE-BASED COMPENSATION IMPLEMENTATION AND IMPACT IN HIGH PERFORMING SCHOOLS IN SHANGHAI, CHINA

BY

PRIYA GOEL LA LONDE

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Educational Organization and Leadership
with a concentration in Educational Administration and Leadership
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2017

Urbana, Illinois

Doctoral Committee:

Associate Professor Anjalé Welton, Chair
Professor Christopher Lubienski, Indiana University
Professor Jennifer Greene
Professor Cameron McCarthy

ABSTRACT

Performance-based compensation is gaining traction in the education systems of countries that lead in the world economy and in comparative international assessments of student achievement. Longstanding and contemporary debates in educational policy and research question the potential of this incentivist policy to improve teaching. Some scholars warn that decades of experimental research have found bonuses yield little or no improvement in various measures of teaching quality and student achievement. Yet, policymakers and performance-pay advocates maintain that financial bonuses will lead to better teaching, student learning, and educational markets. As PBC persists, we remain unaware of how financial bonuses are enacted and shape teaching.

To better understand how financial bonuses shape teaching, recent qualitative research has investigated exactly how performance-based compensation unfolds inside of complex school settings. Building on this emerging scholarship, this research examined the implementation of merit pay and its effects on teaching from the perspectives of teachers at School M1, a high performing elementary school in Shanghai. Those who aspire for high ranks in comparative international assessments and other proxies of global economic leadership are turning to regions such as Shanghai for lessons on "what works." Shanghai thus provided the opportunity to investigate incentivism in a system that has gained global influence yet whose policies and corresponding practices remain understudied. A principal and 20 teachers were interviewed, and a range of policy documents were collected in This qualitative case study. Data were coded according to theory on incentivism, policy enactment, and teacher development. Responses were then further analyzed to determine recurring themes and patterns.

I find that state and national policymakers and the M1 Principal intended for PBC to

improve teaching quality. However, the M1 teacher participants perceived bonuses were intended to augment base pay, compensate teachers for their existing effort, and to provoke teachers to exert additional effort. Teachers had little understanding of merit pay implementation processes. While merit pay did not erode teacher relationships, and norms of respect and deference to authority guided teachers' silence on and acceptance of the policy and shaped their participation in policy enactment. They perceived high expectancy, low instrumentality and valence, and they held mixed views on fairness.

The majority of teachers suggested their sense of motivation and self-efficacy was driven by a sense of personal responsibility for collective good and by public displays of recognition from their peers and superiors. Teachers suggested that social relationships, professional community, and trust were more efficacious ways to get to improved teacher quality. Overall, teachers perceived bonuses neither inspired instructional improvement compelled their low performing peers to leave the school or the profession.

Advocates suggest bonuses motivate educators to behave in desirable ways, incentivize strong teachers to join and stay in the workforce, and force weak teachers to exit. While compelling, this lens ignores how financial incentives operate in the cultural foundations of schools. Additionally, teacher perceptions of bonuses bear directly upon how bonuses are levied in schools. To assess the efficacy of PBC as a school improvement mechanism, this research offers a sorely needed school-level investigation of PBC that explores the perspectives of teachers and is founded in sociocultural lenses on school improvement. In doing so, this research builds on an emerging body of qualitative research that helps assess the prevailing hypothesis that there is a neat relationship between an incentive, teacher practice, and student learning.

Keywords: performance-based compensation, global education policy, Shanghai

ACKNOWLEDGEMENTS

Thank you to the American Education Research Association; University Council of Educational Administration; Institute of International Education; China Confucius Institute; University of Illinois Graduate College; and the University of Illinois Department of Education Policy, Organization, and Leadership for their generous funding and support of my research. Thank you to the Educational Administration and Leadership faculty – Don Hackmann, Mary Herrmann, Linda Sloat, and Anjalé Welton – for a rigorous and rewarding doctoral experience. I also thank all of my UIUC friends and colleagues for their collaboration and friendship.

Thank you to my friend Huriya Jabbar; dissertation committee members Jennifer Greene and Cameron McCarthy; doctoral adviser Christopher Lubienski, and my mentor Anjalé Welton for their invaluable feedback on this research and my professional pursuits. Each of their unique lenses have elevated the quality and significance of my research. They challenged me to think deeply, coached me through through publication and research projects, and fostered my scholarly development. I am forever grateful for their unwavering commitment to my success.

I also thank my family for their support of my research. I am particularly grateful for the incredible generosity and support of my parents, Anita and Om Goel. My niece and son, Rumi and Riyaan, are my future. I wish to remain a model of kindness and perseverance in their lives. My heartfelt gratitude goes to my husband, Robert, for his unwavering patience and love.

Finally, I thank my colleagues at East China Normal University – Li Jiacheng, Lu Keyi, Yang Yang, and Zheng Xue – and the participants in this study for their generosity and compassion. They welcomed me into their culture and their schools; trusted me to well represent their intimate thoughts and dreams; and they pushed me think differently about teaching, learning, and leadership. I am eternally grateful for their kindness and openness.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Since the 1980s, across developing and industrialized nations, reforms in educational policy and practice increasingly feature accountability and incentives. Often referred to as the Global Education Reform Movement (GERM; Sahlberg, 2011), these reforms consist of the standardization of public schooling, private participation in the provision of public education, and the creation of a mass of high-quality teachers, also known as teacher professionalization. The competition, free market underpinnings of these approaches are said to combat longstanding obstacles to school improvement such as ineffective school boards, bureaucratic teacher unions, and inefficient centralized control and funding (Walberg & Bast, 2003.) Advocates purport that through standardization, privatization, and teacher professionalization, innovative schools will rise, poor schools will exit the market, and the quality of educational markets will improve, particularly for underserved students (Chubb & Moe, 1990).

One such reform that is squarely aimed at teachers and teaching is Performance-Based Compensation (PBC). Also known as performance- or merit pay, these financial bonuses are granted to individuals or groups of employees based upon the attainment of predetermined performance criteria. Australian Parliament member Andrew Leigh (2012), defined merit pay as, "instances in which teachers receive temporary or permanent salary increases for being more effective in the classroom" (p. 1). PBC reflects a New Public Management culture of delivery and performativity in school reform (Sahlberg, 2011; Trujillo, 2014). In the context of the teacher workforce, bonuses are said to motivate educators to improve instruction, incentivize strong teachers to join and stay in the workforce, and encourage weak teachers to exit (Hanushek & Lindseth, 2009). Advocates claim individual and group bonuses improve instruction, attrition,

professional interactions, and creativity, and generally motivate teachers to behave in ways that improve teaching outcomes (Clees & Nabors, 1992; Lavy, 2007; The New Teacher Project, 2014).

Performance-Pay: The Nexus of Standardization and Incentivism

Performance-pay is situated at the crossroads of standardization and incentivism. Standardization holds educators accountable for teaching a common curriculum and for students achieving at uniformly high quantitative markers of proficiency. Under this framework, learning objectives, curriculum, and assessment are aligned with one another and proficiency benchmarks are established. Standardization is attractive to policymakers because the combination of core/common curricula, lofty performance targets, frequent and onerous testing of students and teachers, and test-based accountability create the homogenous conditions necessary to improve educational markets in efficient time frames and at low costs (Sahlberg, 2015). Embedded in standardization is a focus on core academic subjects such as literacy/language arts, mathematics, and science and favoring common, scalable, and measurable pedagogical strategies that guarantee teachers will "deliver" the standardized curriculum. State and national annual examinations (i.e., "high-stakes tests") and national learning standards are thus natural features of standardization. The U.S. 2001 Common Core State Standards (CCSS) and Great Britain's 1988 Education Reform Act represent two accounts of first-time statutory policies that created national frameworks for curriculum, and consequently, for standardized assessments (Meyer, Tröhler, Labaree, & Hutt, 2014; Yarovaya, 2015).

In addition to standardization, the current education reform movement exercises accountability for teaching improvement through incentivism. Teachers are rewarded through financial bonuses and career pathways, and sometimes punished, based upon standardized

metrics for preparation, development, and performance. These metrics are embedded in the test-based student proficiency and growth systems enabled by standardization. Policies like PBC are thus viewed as an efficient strategy to improve the teacher labor force, and consequently, create more competitive educational markets (Rivkin, Hanushek, & Kain, 2005). One technique used to determine rewards and punishments, Value-Added Modeling (VAMs), is used in sales and operations. These models use formulas to attempt to exact and ascribe a particular value of a teacher to a student's or a group of students' performance. Existing merit pay programs in the U.S., for example, use VAMs to determine bonuses, promotion, and retention. As a result, incentivism is tightly coupled with teacher and principal promotion, development, and evaluation. The U.S. Race To The Top competition, England's Sure Start, and Ontario, Canada's Literacy and Numeracy are examples of national policies that use performance-pay as an accountability device within an overall scheme of incentivism and standardization.

The Penetration of Performance-Pay In Global Education Policy: The Case of Shanghai

PBC is highly visible in global education policies (GEP) targeted at teacher accountability and improvement. Merit pay has penetrated global education reform rhetoric, recommendations, and policy under the economic premise that bonuses lead to optimal individual performance (Weiner, 1980). Shanghai is one of a few regions that has adopted PBC across an entire state or nation. The case of Shanghai provides an opportunity to fill in the black box of incentives at the school level and rigorously examine how financial bonuses occur through behavioral, social, and cultural mechanisms inside of schools. By studying Shanghai, we can better assess how incentives perform on a large scale and to what extent incentives contribute to improved teaching. Specifically, Shanghai's PBC policies and practices merit analysis for three main reasons: 1) Shanghai's prominence in GEP discourse is based upon

popular rhetoric and thin analyses, 2) the shared Confucian learning and teaching cultural norm is rare, and 3) Shanghai is the central government's focal point for domestic policy goals.

Is It So?: Popular, Thin Rhetoric Needs An Empirical Boost

The Organisation for Economic Cooperation and Development (OECD) and other policy actors in GEP have positioned Shanghai as a world model for education reform. Shanghai has gained international recognition as a high performing education system with strong teaching and learning outcomes. Today, this metropolis of over 20 million people is featured in global education policymaking forums, such as the International Education Roundtable, and in conferences hosted by the Asia Society and OECD. This recent spotlight is due in large part to Shanghai students' superior performance in the Programme of International Assessment (PISA), the OECD's famous comparative international assessment. In 2009 and 2012, Shanghai students ranked first, and in 2015 they ranked in the top ten in reading, science, and math (PISA 2009 Results; PISA 2012 Results; PISA 2015 Results).

Increasingly, high profile comparative assessments, such as PISA, are used to standardize progress and competition across the world (Berliner, 2015). Today, PISA and other comparative international assessments govern the global framing of accountability and the global aims of education (Labaree, 2014; Meyer, Tröhler, Labaree, & Hutt, 2014). Countries that are competing for superiority in PISA and in other proxies of global economic leadership (Meyer, 2013; Sellar & Lingard, 2013a) deem Shanghai an exemplar of 'what works' in schools. Seller and Lingard (2013b) explained how post PISA, the U.S. framed Shanghai as a positive archetype of school reform and a leader in the "global education race":

> Tucker's (2011) edited book, *Surpassing Shanghai*, constitutes a significant intervention into contemporary education reform debates in the US, at a time when US policy-makers are looking outwards for reform ideas. Drawing on research conducted by the OECD and the US National Centre on Education and the Economy, requested by US education

> secretary Arne Duncan following the publication of the PISA 2009 results, the book provides five case studies of top-performing education systems 'in the highest ranks in terms of quality, equity and productivity' (Tucker 2011, 172). These include Shanghai-China, Finland, Japan, Singapore and Canada…The theoretical framework for the study is based on a model of the relationships between economic and educational development, with the goal for nations being the complementary development of high-skill, high-wage knowledge based economies and professional, creative, mass education systems to serve them. (p. 717)

Western policymakers are turning outward to "reference societies" such as Shanghai (Seller & Lingard, 2013a) to identify effective accountability mechanisms in high performing systems. Yet this rhetoric has yet to be substantiated by rigorous, peer-reviewed research. Researchers have examined pedagogical approaches unique to Shanghai and neighboring countries such as lesson study and call and response. Scholars have also disaggregated PISA data and highlighted disproportionately affluent student sample unique to Shanghai students who participated in PISA. However, studies of teacher improvement policy implementation and impact in Shanghai are not available in internationally reputable, peer-reviewed research journals. To move beyond thin accounts of policy and practice and to better understand the source of student's strong performance, rigorous, school-level studies of contemporary policy are sorely needed.

Confucian's Golden Rule: A Shared Learning and Teaching Culture Across Hundreds of Schools

Chinese people share differences in language, cuisine, environment, wealth, and ethnicity. Yet generations and millions of people share a common understanding of and appreciation for Confucius, whose "Golden Rule" reminds the Chinese people of the law of reciprocity. Confucius's teachings undergird the culture of learning and teaching across Shanghai's 1,500 schools. Confucian principles of community, individual accountability for collective success, fairness through equality, long-term visions for success, and harmonious

relationships permeate this high performing education system's philosophy of learning and teaching (Bozionelos, & Wang, 2007; Sellar & Lingard, 2013).

Stemming from an appreciation for teamwork and collective success, authentic teacher-centered improvement activities such as peer observation and team-based lesson planning are hallmarks of teaching culture. The majority of new teachers receive ongoing mentorship from veteran teachers, participate in team-based professional development for at least 50 hours per year, and most school leaders have an average of 15 years of teaching experience (OECD, 2012a; OECD, 2012b). Shanghai thus reveals a rare opportunity to gather teachers' perspectives on financial bonuses in school cultures that weave standardization and incentives with norms of community and equality.

The Country's Shining Star: Fulfilling Domestic Policy Goals Through Shanghai

Shanghai's 1990s *First Class City, First Class Education* policy is grounded in broader goals of leading in the global knowledge economy. The policy calls for education reforms in order to innovate, raise international competitiveness, and enhance educational quality (Lai & Lo, 2007; OECD, 2012a; Tucker, 2011). These reforms include curricular changes such as project-based, research-oriented, and Information Technology learning; reforms to school admissions criteria; and reduced pressure and administration of testing. Still though, annual and rigorous examinations persist as young as third-grade, and student performance is closely coupled with financial bonuses and teacher evaluation.

PBC is a key aspect of compensation reform in Shanghai. Since 2009, teacher salaries are comprised of 70 percent base pay and 30 percent merit pay. Individual and team bonuses are based a range of metrics, such as student performance on benchmark assessments and participation in professional development. Shanghai awards annually a minimum of $2,000 to

master teachers whose students perform well on exit exams (OECD, 2012b).

The Understudy

England, Wales, and Singapore are examples of the few systems that have national bonus pay systems. Empirical investigations of performance-pay implementation and impact that are grounded in teacher lenses and sociocultural theoretical approaches in these systems would also afford a better understanding of how financial bonuses shape teaching. Yet these systems do not share the unique school and system features that belong to the case of Shanghai. Though influential in GEP because of its strong stature in the West, England and Wales are not contemporary exemplars of "what works."

The commonness of Confucian learning and teaching culture also cannot be found in these three countries. Each country boasts a great deal of ethnic, linguistic, and cultural diversity. School choice abounds, and families have a range of options, including public, private, charter, and specialty schools. School type and student community are two prominent factors that make school culture quite distinctive across school sites. Indeed, each of Shanghai's schools hold distinguishable cultural characteristics – there are marked differences by class, migrant status, and location. However, on the whole, the Confucian values that underpin learning and teaching are relatively homogenous across school sites, which enables the potential for a rigorous comparative analysis of performance pay across school sites and districts. Although it is a system that is achieving well, and gaining global prominence, Shanghai remains understudied. provides sorely needed analysis of the penetration and performance of PBC, a globally popular policy device in teacher policy.

Not 'What Works' or 'Does It Work,' How Does It Work?

Yet, research has suggested PBC has not led to, and in some cases has detracted from, school improvement aims. Experimental studies show PBC has done little to substantively improve teaching outcomes, such as effort and instruction (Balch & Springer, 2015). Scant qualitative research suggests financial incentives incite job stress; stifle collegiality; and may not be fully implemented because of resource limitations, capacity weaknesses, and educators' perceptions of a lack of fairness (Rice, Malen, Jackson, & Hoyer, 2015). Nonetheless, the policy is gaining traction. As PBC gains traction globally, we must better understand how incentives operate and the process through which the consequences of performance-pay unfolds in complex school settings. Once we look inside schools and communities of teachers, we can begin to assess the efficacy of merit pay as a mechanism to improve teacher quality.

Rather than hypothesizing what works or whether performance incentives work in Shanghai, this research digs deep into the black box of incentives inside of schools. I ask how performance incentives work from the vantage point of teachers –the main conduits of policy enactment and of teaching improvements. I implore teachers to paint a picture of how performance-pay is enacted within the culture of learning and teaching. I examine their perspectives from sociological and behavioral lenses. These questions and guiding theories are rare but necessary for three key reasons. Firstly, knowing enactment leads to knowing where policy intentions and reality breakdown and that can lead to lessons for better policy design and implementation. Second, school culture is empirically linked to teaching and school improvement. Knowing the cultural conditions in which the policy is embedded helps pinpoints the exact cultural issues (e.g., trust, equality) that are at odds with the policy. Third, teacher perspectives illuminate ways to improve the policy and to improve teaching overall that may

otherwise have gone unnoticed.

Dissertation Overview

In this qualitative case study, I gathered teachers' perspectives on how performance pay was enacted in School M1 in Shanghai and how bonuses shape instructional improvement.

In Chapter 2 I examine empirical and descriptive literature on PBC. The descriptive literature provides an important overview of the recommended, and currently popular, PBC models in P-12 education. I find that, on the whole, experimental studies on the effects of PBC have found bonuses do little to substantively improve student and teacher outcomes. England, Wales, and the U.S. are well pronounced cases in PBC literature. The research reveals several key gaps in the literature that are addressed in this study: 1) how policy actors perceive performance incentives; 2) how PBC operates in high performing schools with complex school cultures; and 3) how policymakers' and school actors' perceptions of incentives vary.

In chapter 3 I explain the theoretical frames that I used to explore the connections between and processes through which incentives, motivation, performance, and improvement unfold in the sociocultural school settings. Experimental research has relied mostly on labor economics' notion that pay has great potential to improve performance. I use theories that get closer at culture, context, and behavior in order to assess the claims that undergird incentivism.

In Chapter 4 I describe the research design that guides this dissertation study. This study utilized purposeful sampling in a qualitative case study (Yin, 2003) of teacher merit pay in School M1, a high performing elementary school in Shanghai. I collected 21 interviews with teachers and administrators in these two schools. Also, I employed a constructivist approach to the collection, translation, interpretation, and analysis of the multilingual interview data.

Chapters 5 highlights key findings on the enactment of PBC policy within a Confucian

school culture context, the divergent aims on PBC, and on teacher perceptions of how bonuses shape teacher practice. In the final chapter of this dissertation I discuss the implications of these findings. I offer implications for policymakers who wish to consider alternative ways to ignite improvement in instruction and student learning, research on incentivism, and implications for school leaders who face the task of building school culture and enacting school improvement in the presence of performance appraisal mandates.

CHAPTER 2

REVIEW OF LITERATURE

In this chapter I review the existing literature on PBC. Descriptive literature has outlined recommended, and currently popular, PBC models in P-12 education. Most experimental empirical research has investigated the effects of PBC and found bonuses do little to substantively improve student and teacher outcomes. Unlike the understudied case of Shanghai, England and Wales and the U.S. are well pronounced cases in PBC literature. Gaps in the literature include research on how school actors perceive performance incentives, as well as the ways in which PBC operates in high performing schools with complex school cultures. I review literature on policy design features, popular policy approaches, and the effects of empirical literature about performance-pay on teaching and learning.

Performance-Based Compensation Program Design Components

Compensation reform includes market-based pay, knowledge and skills-based pay, career ladders, recruitment and retention awards, and pay-for-performance (Springer, 2009). Incentives are inherent across these approaches, though pay-for-performance typically incentivizes student performance in particular. Malen and colleagues (2015) suggested goals, eligibility criteria, award characteristics, and award determinants distinguish PBC programs. These characteristics are featured in PBC program structure, pay allocation, and performance measures (see Figure 1; Odden & Kelley, 1997; OECD, 2009, Podgursky & Springer, 2007a, 2007b; Springer, 2009).

Program structures. PBCS generally favor absolute performance in fixed contracts over relative performance in tournaments. Tournaments award a certain number of top performing individuals a predetermined bonus, whereas fixed contracts award bonuses to all individuals who meet a performance threshold. For example, in a tournament (i.e., relative performance), the top

ten scoring eligible teachers receive 1000 USD, while in a fixed contract (i.e., absolute performance), all teachers who meet the performance threshold receive 1000 USD. In the 1980s, tournament-based PBCS were shown to create competition and disrupt collaborative ethos and are thus often looked upon unfavorably (OECD, 2009). Yet, fixed bonuses are not without their own challenges. Collective bargaining negotiators often advocate for the lowest performance thresholds possible, thereby maximizing the possibility for high threshold pass rates (Heneman, Milanowski & Kimball, 2007). In 2002, for example, England's first year of Performance Management, 97 percent of teachers met the fixed performance threshold (Haynes, Wragg, Wragg, & Chamberlin, 2003a, 2003b). High success rates and uncertainty with costs can cause taxpayers and the government significant, unforeseen financial burdens. Despite this, fixed bonuses are viewed as pragmatic and aligned with teacher culture and thus are preferred.

Incentives are typically organized around individual, group, or hybrid units. The "free-rider" problem (i.e., subpar performers benefit equally from high performers' success) is a grave challenge associated with group bonuses. This is one among many reasons economists suggest individual incentives are the most likely to improve labor market selection and composition of human capital (Lazear, 2003; Podgursky & Springer, 2007). Yet, in the context of education critics warn bonuses for individual teachers are complicated by several phenomena: student performance is not easily attributable to an individual teacher, individual rewards are antithetical to teamwork goals in teaching, the effects of learning are cumulative and build from year to year, and non-school factors greatly affect student learning. As a result, hybrid models are quite popular. In the U.S., for example, in the Milken Family Foundation's Teacher Advancement Program (TAP), student performance on annual state standardized tests determine bonuses for individual teachers and entire schools. A significant challenge associated with hybrid models is

that data systems must aggregate and disaggregate data in order to link student performance to both individual and groups of teachers.

Pay allocation. The OECD (2009) offered several illustrative examples of how bonus size and distribution vary greatly. Bonus size largely depends on the number of individuals expected to receive bonuses. Two popular solutions are to give smaller rewards to more teachers or to give larger rewards to fewer teachers. Over time, the sizes of bonuses have increased due to greater interest in testing whether larger incentives lead to improved teacher performance (Springer, 2009). While size and thresholds vary, distribution typically occurs at the end of the school year, in between school years, or at the start of a new school year because often student performance tests are administered late in the school year. Research suggests that with regular and timely bonus allocations, system gaming behaviors decrease and perceptions of bonuses improve (OECD, 2009; Vroom, 1964).

Performance measures. The measures used to determine performance bonuses are perhaps the most widely disputed issue in PBC policy. Some advocate for input-based rewards, wherein teachers earn bonuses for participation in instructional improvement activities such as professional development, peer mentorship, teacher leadership, and educational advancement (Odden & Wallace, 2002). Often referred to as knowledge- and skill-based pay, these measures are said to mitigate issues of 'teaching to the test' and disproportionate focus on a certain subset of students and motivate individuals to focus on self-improvement. Some scholars claim this approach is flawed because observable teacher characteristics (e.g., educational attainment, certification) explain little of the difference in student achievement (Goldhaber, 2002; Berliner & Glass, 2014).

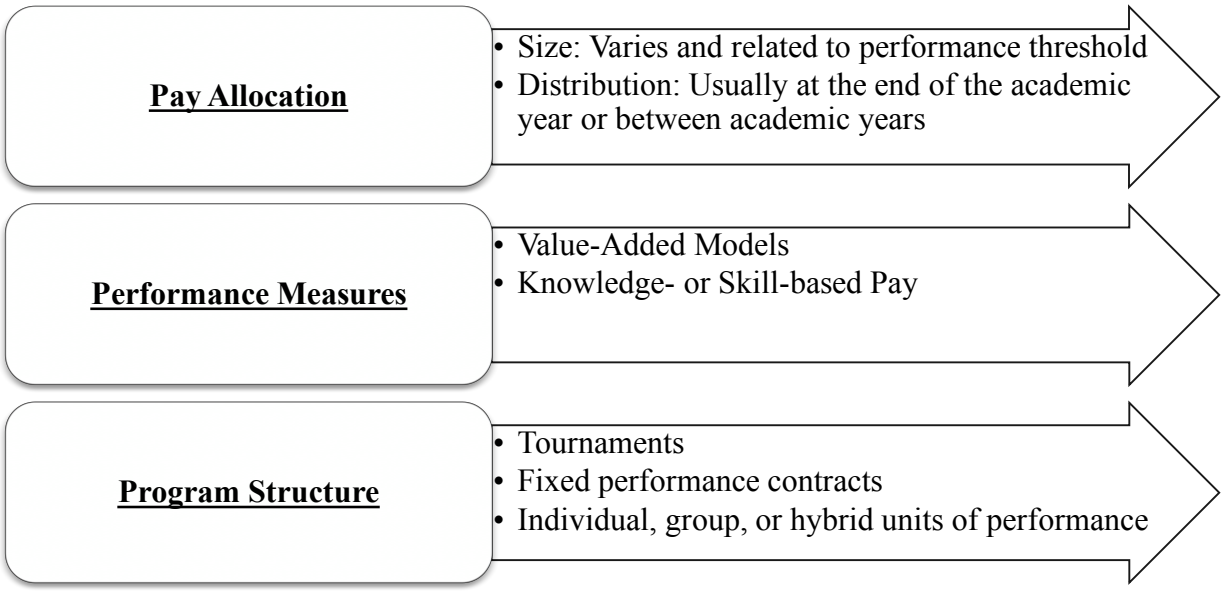
Measuring teacher quality through student performance. Currently, many PBC policies

mandate the use of student performance measures to partially or fully evaluate teachers and/or award merit pay (OECD, 2009; Teacher Incentive Fund, 2015). These measures include student academic performance, retention, attendance, and advancement (i.e. credit hours or passing grades). In the U.S. most states use student achievement data in Value-Added Models (VAMs) or growth models (Student Growth Percentiles, SGPs) to evaluate teachers (Race To The Top, 2011), and in some cases, to award teacher merit pay. Both models are designed to determine the portion of student performance on standardized assessments, or the "value-added" score, that is attributable to a teacher (Hallinger, Heck, & Murphy, 2014). VAMs typically control for student prior history and student- and school-level variables in more complex ways than SGPs (Amrein-Beardsley, Polasky, & Holloway-Libell, 2016). Because both models typically use largescale standardized tests to determine the teacher value-added score, teachers of non-tested grades and subject areas are typically ineligible for individual, teacher-level value-added scores. To evaluate the performance of teachers in non-tested grades and subjects, states are using teacher-developed student learning objectives (SLOs) and benchmark assessments (e.g., Measures of Academic Progress) to measure student growth and determine the teacher value-added score. However, SLOs and benchmark assessments are not widely recognized as "objective" determinants of teacher performance, especially as compared to the largescale standardized assessments, VAMs, and SGPs.

Scholars warn that test-based accountability is a misguided, inappropriate measure to evaluate and reward teachers for several reasons. Longstanding research suggests student background factors including race, class, and parents' educational attainment more closely determine student achievement than teaching (Ballou & Podgursky, 2003, Berliner & Glass, 2014; Lubienski & Lubienski, 2014; Rothstein, 2004). Some scholars have found in-school

factors, including teaching, only explain 14 percent of student performance (Berliner & Glass, 2014). Also, teachers, school leaders, itinerant staff, and coaches, who may or may not be involved in tested grades or subjects, contribute to student performance. Additionally, standardized and interim benchmark test items are biased, particularly against emergent bilingual students, students of color, and students from low-income families. Proficiency models implicitly encourage teachers to focus on the "bubble kids" in order to move as many students as possible across proficiency thresholds, and growth models implicitly encourage focus on the "low kids" in order to demonstrate the most amount of growth from as many students as possible (Jennings, 2012). In some cases, when students show inadequate proficiency or growth, punitive consequences are ascribed to teachers, such as public shaming, employment probation, or termination (Berliner & Glass, 2014; Orland, 2015). Despite these critiques, student performance remains a central measure of PBC. In the next section I review current policy approaches to and research on the effects of performance incentives.

Figure 1. Performance-Based Compensation Programs Design Components



Policy Approaches to Performance-Based Compensation

Distinctive approaches to PBC are readily visible in the U.S. and parts of the U.K. U.S.

PBC is varied and shaped by multiple levels policy, whereas in England and Wales, as of 2013, all public school teachers are subject to performance-pay. These countries are the first to tie teacher salary progressions entirely to performance metrics.

England and Wales: A national performance-pay system. In 2000, the England and Wales Department for Education and Employment (DfEE) piloted a two-pronged Performance Threshold program – Performance Management (PM) and Threshold Assessment (TA; Haynes et al., 2003a, 2003b). PM was a teacher appraisal policy, while TA linked teacher performance with a 2000 Sterling Pound annual bonus and progression from the standard (i.e., Main Pay Scale) to the higher pay scale (i.e., Upper Pay Scale). Policymakers unveiled TA first in an effort to immediately award experienced teachers, gain the support of unions and professional associations during a critical elections period, improve teacher recruitment and retention, and modernize the profession. However, educators perceived the policy as "something for something" and an effort to link employment to performance (Mahony, Hextall, & Menter, 2002). Although the government articulated an inclusive policymaking process, unions and associations were not involved in the policy design. Team leaders (i.e., head teachers) and external Threshold Assessors (i.e., inspectors) approved 97 percent of 4,753 teachers who applied for bonuses. Thereafter, government agencies published an appraisal model to carry out PM. Principals assigned a team leader to observe classrooms; hold information discussions; and plan, monitor, and review objectives with all teachers regarding about five standards: knowledge and understanding, teaching and assessment, pupil progress, wider professional effectiveness, and professional characteristics (Haynes et al. 2003a, 2003b; OECD, 2009).

In 2006, the Performance Threshold system became statutory. It was revised to more closely align with new professional standards and career stages and aimed to increase

transparency and efficiency in distributing rewards (Evans, 2011). All teachers were required to demonstrate proficiency in two stages below the threshold (i.e. Qualified and Core), and they had the option to apply to cross the threshold and earn higher salary in three additional stages (i.e., Excellent and Advanced). Team leaders and Threshold Assessors determined whether an individual crossed the threshold through quantitative markers of student achievement.

In the 2010 White Paper, *The Importance of Teaching*, the DfEE announced plans to reduce the number of standards, align standards more closely with elements of teaching, and more clearly "identify and deal with unsatisfactory performance" (DfEE, 2010 in Evans, 2011). As of the 2013-2014 school year, Performance Threshold now includes four key features: 1) performance appraisals determine salary progressions for teachers on the Main and Upper Pay Scales; 2) educational attainment or years of service no longer determine salary progression; 3) salary progression is no longer bound to points or fixed increases, but there are minimum and maximum salaries in each scale; and 4) senior colleagues determine teacher performance through the teacher standards and progress on the teacher's previous year objectives as outlined in the School Development Plan (i.e., school improvement plan; Coughlan, 2014; Marsden, 2014; Ratcliffe, 2013; Walker, 2013). In 2013, this policy affected the approximately 37,000 full-time teachers in Wales and 451,000 full-time teachers in England who worked in state-run schools (Coughlan, 2014; EWC Annual Statistics Digest, 2014). Since 2010, both regions have increased the number of teachers with graduate degrees and with Qualified Teacher Status, but higher proportions of teachers have exited each year.

United States: Merit pay as a way to the top. In contrast, the U.S. has a much longer history of experimentation with PBC. The earliest experiments, which were between the World Wars and again in the 1980s, (Murnane & Cohen, 1986) were used to create alternatives to the

single salary schedule. Leigh (2012) warns that political support for PBC does translate neatly to political sustainability. In the U.S., 75 percent of merit pay programs in 1983 were defunct ten years later. Since 2006, the U.S. Department of Education's Teacher Incentive Fund grants (TIF) has funded 131 federally funded programs in 2000 schools across 36 states and Washington D.C. to "use of performance-based compensation, and other human capital strategies that enhance and sustain performance-based compensation, in order to increase students' access to effective educators in high-need schools, and to expand the array of promising approaches that can help these educators and other personnel succeed" (Teacher Incentive Fund, 2015).

PBC is also a cornerstone of the $4.35 billion Race to the Top (RttT) competition. All 19 RttT states currently have or in the near future will incorporate PBC into compensation policies in participating school districts (United States Department of Education, 2015). Minnesota, Florida, and Texas together have allocated more than $550 million to performance pay programs (Springer & Gardner, 2010). In addition to TIF and RttT funds, states and districts have funded PBC through bond and tax increases. Some states keep tight control over district's compensation reforms through state-level per pupil funding. For example, the state of Washington supplies 70 percent of every district's funding. The state allocates funding according to its state teacher salary schedule and stipulates that states can add on to salaries through one-year contracts that stipulate additional work time, additional work responsibilities, and/or incentives.

Brodsky and colleagues (2010) examined programs in three states – Minnesota, Florida, and New Mexico, and in three districts – Denver Public Schools, Austin Independent School District, and Houston Independent School District – and found variance in program structure, performance measures, and pay allocation (see Table 3). Founded in 1999 and the oldest of the six program, Denver's veteran teachers who opt to participate in PBC remain in the performance

pay program for his/her entire career in the district. However, new teachers must participate in the program. In Houston and Florida, student performance on annual standardized tests solely determines incentive pay; whereas in Minnesota, New Mexico, and Denver, student performance, teacher knowledge/skills, and supervisor evaluations jointly determine performance pay. Austin and Denver offer incentives to teachers who work in hard-to-teach/staff subjects and schools. Intermediary organizations (e.g., advocacy organizations, foundations, think tanks) have participated in the conceptual, programmatic, and fiscal development of each program. Bonus allocations vary greatly from $1000 to $5000 per teacher per school year, which represents between two to five percent of the average U.S. teacher's annual salary.

Policymakers have not found a neat solution to offer performance-pay to teachers who teach in untested grades and/or subjects. The fine arts are a prominent example. Elpus (2011) cites three key problems with PBC for arts teachers: a) the "high quality teacher" definition is vague, especially as "high quality teacher" is increasingly determined by quantitative measures of student performance; b) evaluation of the "high quality" arts teacher in the arts is also vague, and c) there are inequities in access to merit pay. For example, in Washington D.C. and Denver school districts, arts teachers are not privy to the same amount of bonus pay as teachers of reading and math (Elpus, 2011).

Effects of Performance Incentives

Several decades of research have examined the effects of PBC on student and teacher outcomes. Experimental data on the effects of PBC are mixed, and overall, merit pay claims rest upon weak empirical support. Schools adopt performance incentives primarily through group-based programs. In rare instances, incentives have yielded slight, positive student and teacher outcomes in the U.S., Israel, India, and other OECD nations. Yet, longer-term effects are

unknown because these programs have not persisted over time.

Experimental Evidence on Student Outcomes

Famed author on human motivation, Daniel Pink, summarized the performance of financial bonuses across disciplines:

> Here's what science knows: 1) Those 20th century rewards, those motivators we think are the natural part of business do work, but only in a surprisingly narrow band of circumstances; 2) those if-then rewards often destroy creativity; 3) the secret to high performance isn't rewards and punishments but that unseen intrinsic drive – the drive to do things for their own sake, the drive to do things cuz [sic] they matter (TedGlobal, July 2009)

The bulk of experimental research has examined U.S. program outcomes; but some data have looked at program effects in other regions such as Kenya, India, and Israel (see Table 1). Mathematica Policy Institute's evaluation of U.S. 2010 Teacher Incentive Fund (TIF) districts' PBCS found that just half of all districts implemented all four required program components, 65 percent of districts reported sustainability was a challenge, and many districts reported PBC was organized in that did not yield successful implementation (Chiang, Wellington, Hallgren, Speroni, Herrmann, Glazerman, & Constantine, 2015). Teachers who were part of the ten districts that participated in a sub-randomized control trial reported limited understanding of the program, their eligibility, and reward potential. In these ten districts, students in the treatment group showed a slightly positive improvement in reading but no statistically significant improvement in math.

Studies of PBCS in New York City and in Austin, Texas have found either slight or no effect on student performance (Balch & Springer, 2015; Marsh, Springer, McCaffrey, Yuan, & Epstein, 2011). One study of the 2007 New York City group-based PBCS found, "in schools where smaller groups of teachers were responsible for instructing tested students, the program led to small but significant increases in student achievement" (Goodman & Turner, 2013, p.

410). The authors suggested linking pay to school-wide performance goals, the presence of a parallel accountability system that sanctioned schools that did not meet the same goals that were used to determined bonuses, and teachers' lack of understanding about the PBCS limited program effectiveness. Sojourner and colleagues' (2014) analysis of Minnesota's Quality Compensation (QComp) program found a slight increase in achievement for students whose teachers were less experienced. In their comparison of a traditional and performance-pay compensation systems, Eberts and colleagues (2002) found performance pay increased high school students' course completion, did not affect grade point average, and negatively affected student attendance and passing rates. Some scholars have experimental studies' methodological weaknesses. Lavy (2007) discounted findings on the PBCS in North Carolina, Kentucky, and Texas for not having control groups, weak comparison groups, and voluntary teacher participation. In his review of PBC literature, Leigh (2012) called for longitudinal studies of well-established programs that use a mix of methods.

Survey Data On Teacher Outcomes

Survey research have provided insights into teacher support for merit pay. Matthes and Tollerud's (1990) survey found 247 teachers generally did not support career ladder and merit pay concepts. Women were particularly opposed to peer evaluation in bonus programs, and elementary teachers and teachers with less postsecondary education responded less favorably to incentives and competition-based compensation. National survey data from the U.S. in 1987 suggested private school teachers supported performance pay more than public school teachers (Podgursky and Springer, 2007a). Muralidharan and Sundararaman's (2011) survey found the majority of U.S. teachers (86 percent) had favorable views toward merit pay, while most U.K. teachers disagreed with the principles of merit pay and using student performance measures to

allocate merit pay.

Goldhaber and colleagues (2008) analysis of 1999-2000 School and Staffing Survey data found that districts in states with stronger accountability policies, Right to Work laws, and no collective bargaining were more likely to have PBC programs. They also found teacher total income in merit pay districts was about $621 more than teacher total income in nonmerit pay districts. Approximately 55 percent of surveyed teachers favored merit pay, 12 percent of the schools surveyed used PBC, and ten percent of teachers in these schools received merit pay. Teachers in cities and teachers in private schools were more supportive of merit pay than teachers who worked elsewhere or worked in public schools. Teachers with more experience, women, Whites, those in rural communities, those who believed evaluations in one's schools were not fair were more likely to oppose merit pay. Also, teacher support was not related to school size, student socioeconomic status, or teacher base salary.

In their analysis of 2006 Washington State Teacher Compensation survey data, Goldhaber and colleagues (2011) found teacher self-interest and sense of professional norms of collegiality mediated perceptions of compensation reform. Teachers who expressed high trust and respect for their colleagues, females, teachers with many years of experience, Whites, elementary school teachers, and teachers who identified as members of unions expressed low support for PBC. Teachers of students from low-income families supported merit pay more than other teachers, math and science teachers supported subject-area bonuses, and the mean "fair" amount for a subject-area bonus was $2,331. The authors concluded, "competitive rewards more so than other compensating differentials may be at odds with strongly held professional norms in teaching;" (p. 454) and to reduce the "hard blow at the egalitarian ethos of the profession," (p. 460) districts should allow voluntary participation for veteran teachers and mandate new teacher

participation in PBCS (i.e., the Denver, Colorado approach). In their analysis of several longitudinal studies, Podgursky and Springer (2007b) concluded, "a growing body of research points to large but idiosyncratic teacher effects" (p. 570). Like other scholars, they reported merit pay has neither provoked modifications to teacher instruction nor increased teacher effort (Matthes & Tollerud, 1990; Elliot & Hout, 2011).

Case Studies

Case studies have explored the processes and character of PBC in schools. Though few, these studies offer rich insight into how PBC shapes teachers' perceptions of pay, school culture, and organizational improvement. Kelley and colleagues (2000) examined the organizational structure of, and educators' perspectives, on the Maryland and Kentucky PBCSs. Mandated by the State Supreme Court, Kentucky's PBCS made clear how performance and merit were tied together, and teachers had discretion over how to spend their money. In contrast, the Maryland system had weak, vague sanctions, and schools were required to spend group bonuses entirely on school improvement. About 40 percent of Kentucky schools received rewards, whereas about 10 percent of Maryland schools received rewards. The 1996-1997 interview and survey data showed Kentucky teachers understood their state's accountability system and PBC policies more than Maryland teachers. The Kentucky program yielded teaching improvements, despite negative reactions from principals and teachers. While the Maryland PBCS did not generate strong teacher and school improvement, principals reported more flexibility and comparatively less stress and pressure. The authors concluded a key issue to be explored was, "How much accountability is enough to strengthen teacher collaboration while not undermining the personal well-being of teachers?" (p. 193).

Studies of Maryland's Prince George's County Public Schools (PGCPS), the 18th largest

district in the U.S. with 130,000 diverse students, have shed further light into how PBC unfolds in schools. The district selected 42 high needs schools to participate in a program that used student growth (50 percent of total score for teachers and 60 percent for principals), skills/knowledge growth, and hard-to staff/teach areas to determine merit pay (Malen et al., 2015). In an analysis of interviews with 144 district personnel and 24 leaders, Rice and colleagues (2012) found that as leaders and teachers became increasingly overwhelmed with program demands, their support waned, stress increased, and they gained concern with the "legitimacy and helpfulness of the evaluations and the fairness of the evaluation process" (p. 923). Rice and colleagues (2015) also found issues of fairness abounded. Teachers expected maximum award amounts to be attainable, performance measures to be accurate and reliable, and eligibility restrictions justified, and merit pay allocation rules clear from the outset.

The 2006-2010 PBCS in 15 schools in Austin, Texas, also used multiple metrics to determine merit pay. Performance measures included veteran-to-novice teacher mentorship, classroom level student growth assessment, and schoolwide student growth assessment through statewide peer comparisons. Mentor teachers received $5,500 for mentorship and up to $4,000 in performance-based pay. Teachers received $1,000 to $3,000 for achieving SLOs and $2,000 to $4,000 for schoolwide growth, and principals received $4,000 to $8,000 for campus-wide growth. Austin school district leaders Lussier and Forgione (2010), reported that an internal evaluation found 67 percent of teachers reported the program to be fair and supported student growth measures, and 89 percent of teachers deemed student growth on state standardized tests moderately or highly important. The program was cut because of financial constraints and the Board's position that there was a "lack of compelling evidence that [the program] was having a meaningful impact on campuses," and there was a "lack of national data that incentive-based

programs have had a sustainable impact on school and district performance" (p. 240).

Research on the Kentucky, Maryland, and Texas programs echoed challenges with organizational issues and capacity development, two pronounced issues in PBC implementation literature (Kelley, Conley, & Kimball, 2000; Raham, 2000). Malen and colleagues (2015) discussed how these grave challenges largely shape "partial and symbolic" implementation:

> As prized resources such as time, energy, and talent get dispersed across a growing number of distinct programs, capacity is diluted. In short, new policies and programs can make organizations and the individuals who work in them more fragmented and frenetic rather than more effective and efficient. Under these conditions, even when individuals exert extraordinary effort, programs may wind up being only partially or symbolically implemented. (p. 168)

As PBC scales up, capacity, sustainability, and perceptions of fairness remain key concerns.

England and Wales. England and Wales adopted system-wide merit pay programs that tie public school teacher salary entirely to performance metrics (Haynes et al., 2003a, 2003b). Performance appraisals determine salary progressions for teachers, educational attainment or years of service do not determine salary progression, salary progression is not bound to fixed increases, and senior colleagues determine teacher performance through standards and growth comparisons (Evans, 2011; Marsden, 2014; Ratcliffe, 2013; Walker, 2013). In its first year in 2013, this policy affected the approximately 37,000 full-time teachers in Wales and 451,000 full-time teachers in England who worked in state-run schools, and 97 percent of teachers who applied for bonuses were approved (Coughlan, 2014; EWC Annual Statistics Digest, 2014). Both countries have increased the number of teachers with graduate degrees and with Qualified Teacher Status, but higher proportions of teachers have exited each year (Coughlan, 2014).

Scholars reported that education authorities in these countries saw PBC as a vehicle to award experienced teachers, gain the support of unions and professional associations during a critical elections period, improve teacher recruitment and retention, and modernize the

profession. Educators saw the policy as "something for something" and an effort to link employment to performance (Mahony et al., 2002). Haynes and colleagues' (2003b) reported that teachers and headteachers, who are the Threshold Assessors (TA; i.e., the person who determined merit pay decisions), were concerned about fairness, legitimacy, and capacity for strong implementation given that they lacked funding, training, and agreed upon criteria. Haynes and colleagues (2003a) also found the majority of teachers felt like "victims of bias and management bullying" (p. 43) and were concerned with the TAs' power, fairness, and inconsistency with evaluations and communication. Teachers interpreted TAs' behavior as a signal that TAs wanted "unsuccessful" teachers to leave the school. These perceptions appeared to have influenced underperforming teachers' employment decisions, as 29 percent intended to work in a different school, 50 percent had already left their schools, and 33 percent planned to leave the teaching profession. In a separate study, Mahony and colleagues (2002) found TAs' views of PBC were positive, but teachers' views varied but both groups were concerned about implementation and sustainability. A union affiliate predicted, "schools will only have the cash to allow fewer than one in two teachers who have passed the threshold to progress further up the pay scale" (TES, 14 December 2001 in Mahony et al., 2002). Storey's (2000) analysis of over 4,000 written responses to the government's PBC policy announcement from unions, policymakers, teachers, parents, and advocacy groups found that stakeholders feared and distrusted the PBCS. Storey (2000) explained why an absence of trust bears negative consequences for teacher improvement:

> Trust is viewed as a prime conditioning factor in the literature relating to motivation. It is seen as a form of `social capital' a crucial organizational resource or capability, which can give one organization an advantage over others, even those which ostensibly have similar or even the same pay scheme. Trust can be seen as part of the psychological contract. It is part of the sense of fairness at work, which is a mediating factor and which can allow or disable all types of human resource initiatives (p. 521).

Case studies of merit pay tend to divulge several key findings. Self-interest and complex, multidimensional conceptions of fairness and justice may contribute to the feasibility of opt-in merit-pay schemes. Qualitative examinations of well-established PBCS are needed in order to assess whether and how incentive pay is altering the teaching workforce, a central tenet of the policy (Rice et al., 2015). Furthermore, instruments that identify measures of teacher cognitive skills are needed to better understand the relationship between performance pay, teacher skills, and student achievement (Cissel, 2010). Similarly, Neal (2009) suggested proxies of teacher quality must be clearly defined in order to reward teacher quality and ensure bonuses are given to "the types of achievement that the system is intended to foster," (p. 155).

Discussion

The empirical literature tends toward several key findings. Schools have adopted PBC primarily through group-based incentive programs. There are rare instances in which incentives have yielded slight, positive student and teacher outcomes in the U.S., Israel, India, and other OECD nations, but because these programs have not persisted over time, research has not studied the longer-term effects. Contemporary case studies suggest self-interest and complex, multidimensional conceptions of fairness and justice may contribute to the feasibility of opt-in merit-pay schemes.

Survey findings on teacher support for merit pay are contradictory, especially on the patterns of support from teachers work in schools with high populations of students from low-income families. Teachers seem to favor group-based incentive plans, but we know little about what aspects of and under what conditions incentives are attractive. Overall, teachers respond unfavorably to PBC and suggest that programs that base bonuses exclusively, or mostly, upon student performance measures negatively affect collaboration, trust, respect, and capacity to

learn. Leigh (2012) reviewed teacher perceptions as follows:

> First, teachers have real concerns about the impact of merit pay, particularly on the cooperative spirit in the school. Second, there is significant heterogeneity in attitudes, with new hires and minority teachers substantially more inclined to support merit pay. Third, the kind of merit-pay plan that receives least support is one that uses test scores… On average, the kinds of test-based merit-pay schemes generally favored by economists receive little support among the teaching profession. However, there is enough heterogeneity in teacher attitudes that opt-in merit-pay schemes may be feasible. (p. 20-26)

Research has mostly focused on student outcomes in Western countries. We know little about incentives in high performing education systems are influential in global education accountability and reform, such as Shanghai. Longitudinal studies of large incentive programs offer deeper, more nuanced understandings of the effects of incentives on students, teachers, and schools. In terms of teachers and teaching and school leaders and leadership, rigorous research of new and existing programs must examine how bonus pay affects professional learning communities; instruction, curriculum, and assessment decisions; and the professional behavior of teacher and leaders. In doing so, research can begin to describe whether and how incentive pay is altering the teaching workforce, a central tenet of the policy (Rice et al., 2015). Cissel (2010) suggested future research identify measures of teacher cognitive skills to better understand the relationship between performance pay, teacher skills, and student achievement. Similarly, Neal (2009) suggested proxies of teacher quality must be clearly defined in order to reward teacher quality, and PBC programs align measures and procedures that ensure bonuses are given to "the types of achievement that the system is intended to foster," (p. 155).

CHAPTER 3

THEORETICAL FRAMEWORK

In this chapter, I describe the theoretical frames used in this study. Previous studies have relied primarily on concepts from labor economics, such as agency theory, to study incentivism. This approach has been unable to capture *how* incentives operate and the *process* through which their consequences unfold in complex school settings. These foci are key to judging the strength of the bedrock assumption of performance incentives: bonuses lead to improved motivation, teacher performance, student performance, and school markets. Because incentivism works through the perceptions of school leaders and teachers and is shaped by school culture, we need theoretical tools that place culture and behavior in the forefront. Therefore, I integrate insights from behavioral psychology, such as expectancy and goal attainment theory, with insights from school culture and Confucianism to understand how bonus policies unfolded in the School M1. Also, I use a strategic-relational view on the role of ideas (Verger, 2014) and policy enactment (Ball, Maguire, & Braun, 2012) to examine the policy structure and how it unfolded.

Confucian School Culture

Schooling – and thereby teaching and learning – is a social, cultural, and a political enterprise. Policies featured in standardization and incentivism, that are targeted at teacher improvement, are situated in the social, cultural, and political underpinnings of schools (Bryk & Schneider, 2002). To examine the prevailing hypothesis – there is a neat, positive relationship between a financial incentive, teaching, and student achievement – it is imperative to view performance-based compensation as deeply embedded in the fabric of school culture. The culture of School M1 and other Shanghai schools represents a confluence of teaching and learning and national culture. Theoretical tools on school and Confucian culture provide the multifocal frame

necessary to understand the enactment of incentivism in School M1.

The culture of a school is marked by the values and beliefs held by school members, traditions that persist over time (Higgins-D'Alessandro & Sadh 1997). Scholars have suggested that the core values and beliefs of a school culture are exemplified by the rules, celebrations, traditions, and routines of members of the school community (Robbiens & Alvy, 2004). Gruenert and Whitaker (2015) added that culture is "both a survival mechanism and a framework for solving problems" (p. 4). The earliest sociological thinking suggested culture "is not a problem that needs to be solved, but rather a framework that a group can use to solve problems; it is how we learn to survive, one generation passing down what it has learned to the next" (p. 6). Schein (1992) saw culture as a set of unwritten rules that defined a particular group, and Geertz (1973) saw culture as an interpretative science. Trice and Beyer (1993) explained cultural innovation as more difficult than cultural maintenance. Waller (1932, in Deal & Peterson, 2016) underscored the tension of tradition and innovation:

> Schools have a culture that is definitely their own. There are, in the school, complex rituals of personal relationships, a set of folkways, mores, and irrational sanctions, a moral code based upon them. There are games, which are sublimated wars, teams, and an elaborate set of ceremonies concerning them. There are traditions and traditionalists waging their world-old battle against innovators" (p. 7)

Hofstede's (1983, 1991, 2001) research has examined how workplace values are influenced by culture. His research in 76 countries found compensation systems were a function of six dimensions of national culture: power distance, individualism, masculinity, uncertainty, long term orientation, and indulgence. Also referred to as Confucian dynamism, long-term orientation is marked by persistence, perseverance, ordering relationships by status, the possibility of having many truths, thrift, and having a sense of shame (Hofstede & Minkov, 2010). Bozionelos and Wang (2007) explain long-term orientation as "the extent to which a

particular society espouses the values introduced by the moral philosopher Confucius who lived around the 5thc century BC." In their research on identity in the Chinese workplace, Bozionelos and Wang (2007) found China placed significantly higher than the Anglo-Saxon countries on Hofstede's dimensions of collectivism and long-term orientation. Their research underscored the influence of Confucian values on Chinese work culture:

> The teachings of Confucius have largely influenced the shaping of the Chinese culture (Tung, 1996). One of the pivotal Confucian principles, which shares common ground with collectivism, pertains to the importance of the group and points to the fact that individuals must submit themselves to the needs of the group. Hence, the maintenance of cohesion and harmonious relationships within the work group must be a dominant concern for the Chinese in the workplace. (p. 286)

Confucianism holds a very particular view on fairness: "Equality for all within the same group, and equity for those outside the group. The latter is reflected by a Confucian aphorism saying that there is no worry about scarcity but unevenness (Huang, 1997)" (Fong & Shaffer, 2003, p. 564). Goh and colleagues (2009) suggested East Asia Confucianism and communitarianism are taught through formal education, stories, and religious practices. They characterized teachers with Confucian values as "sociocentric, holistic, allocentric, ensembled, constitutive, contextualist, high-context, connected and relational, conservatist, self-transcedent..., interconnected, and depend on each other for self-definition" (p. 260). It is also well documented that Confucian cultures have historically held teachers and schools especially accountable for student learning (Williams & Engel, 2012).

Policy Enactment

Policy enactment offers a framework to illuminate the character of the Confucian context of school culture in School M1 and the ways in which teacher accountability policies, such as Performance-Based Compensation, interact with this unique culture. Unlike policy implementation, policy enactment focuses squarely on the discursive processes, interpretations,

and decoding of policy that is "done by and done to teachers" (Ball, Maguire, & Braun, 2012, p. 3). Ball and colleagues (2012) explain enactments are "collective and collaborative…in the interaction and inter-connection between diverse actors, texts, talk, technology and objects (artefacts) which constitute ongoing responses to policy…in our schools [policy] is always a process of 'becoming', changing from the outside in and the inside out" (p. 4). To make sense of how teachers perceive and take up PBC in School M1, enactment allows a look at the interpretation and translation of actors at multiple levels. By focusing explicitly on context, enactment calls for an explicit focus on the school-specific factors, micro-institutional variables, and the external environments that are "situated, material, professional, and external dimensions" of policy actions (Braun, Ball, Maguire, & Hoskins, 2011).

Strategic-Relational View on the Role of Ideas

The key idea in the adoption of PBC policy is the economic, rational conceptualization of incentivism, which suggests financial bonuses lead to optimal performance. Verger's (2014) *strategic-relational* view of the role of ideas in policy adoption in GEP is a useful tool to study this idea and how it is operationalized the stated and perceived aims of PBC. Verger (2014) asserted a focus on the multifaceted role of ideas in GEP affords a complex interpretation of policy adoption:

> …Paying closer attention to the multiple roles of ideas in GEP processes and, specifically, to their application in dynamics of policy adoption can contribute to unraveling the macro vs. the micro, convergence vs. divergence, local vs. global and related dichotomies that so often stretch the globalization debate in comparative education (cf. Chisholm, 2012; Robertson, 2012)…[Ideas] need to be understood as constitutive of broader structures and institutions…with causal powers over actors' preferences and decisions. In other words, ideas can operate as both structure and agency, action and condition. (p. 24)

Departing from a constructivist view that focuses on the action and condition aspects of ideas, Verger's *strategic-relational* approach calls for a "dialectical understanding of the

relationship between ideas, institutional change, and strategically selective contexts" (p. 20). The interweaving of ideas, institutions, and contexts illuminates the adoption patterns that occur when policy advocates weave core policy ideas into existing institutional frameworks and prevailing policy and political norms. One *strategic selectivity* that is particularly relevant to adoption of incentivist policies is government ideology. Verger explains, "policy-makers are more inclined to adopt policy solutions that fit within the ideology of the government for which they work or that they represent…Governmental ideologies represent a key filter when it comes to adopting OECD recommendations on educational policy" (p. 21). This influence of ideas is also embedded in Verger, Lubienski, and Steiner-Khamsi's (2016) characterization of the global education industry (GEI) as a social field of networks, cognitive frames, and institutions:

> Networks refer to the more or less stable sets of relationships between political, social and/or economic actors, which work as channels of influence and, in many cases, as more or less formal governance mechanisms. Cognitive frames refer to the types of ideas that social, political and economic actors mobilize to advance their vision of societal problems and preferred solutions. And institutions can be broadly defined as sets of rules, norms and procedures in which actors develop their economic activity and their political strategies (p. 14)

Cognitive frames mobilize policy actors to advance their given policy solutions within institutional and regulatory norms and among networks of a variety of policy actors.

Incentivism: What's The Big Idea?

The cognitive frame of incentivism is readily apparent in the social field of GEP. PBC has been advanced under the persuasion that it meets the demands of globalization, competition, and innovation, and technology advances that enable measurement of individual performance (Xiu & Gunderson, 2013). This labor economics principle of agency suggests monetary, explicit incentives, rather than "fuzzy," social extrinsic motivators (e.g., fear of termination, popularity, respect), motivate individuals to expend their maximum effort and behave in ways that lead to

optimal performance (Kreps, 1997; Weiner, 1980). The policy has been widely adopted in human resources practices in the private sector. Most of the Fortune 1000 firms use PBC (Luthans & Stajkovic, 1999) in forms such as profit sharing, efficiency wages, and firm size-based pay to employees who demonstrate strong sales and leadership (Baker, Jenson, & Murphy, 1988). 'Best practice' suggests PBC be derived through a combination of decentralized managerial wage discretion and a competitive production environment (Xu, 2000). Also, economists argue rather than focusing on the "irrelevant and uneconomic" relationship between incentives and motivation and orchestrators must consider the level (i.e., the total cost of a package for an employer or the total value of a package to an employee), functional form (i.e., the performance incentives), and composition (i.e., the relative amounts of the discrete components of the package such as cash, retirement benefits, travel reimbursement) of PBC design (Baker et al., 1988; Earn, 1982). This view asserts that functional form is particularly important because it bears directly upon the level at which employees value packages (Baker et al., 1998).

In GEP, incentivist ideas have been orchestrated among policy actors at multiple levels toward the adoption of vouchers and parent trigger laws in the US (Lubienski, Brewer, La Londe, 2015). Some scholars have warned that the result of the competition of ideas and interests in democratic arenas and institutions may result in the privatization of public policymaking (Lubienski et al., 2015), as evidenced by idea orchestration toward vouchers, parent trigger laws, and charter schools in, for example, the state of Washington (see Au & Lubienski, 2016). Similar to vouchers and parent trigger laws, the orchestration of PBC has invoked ideas around market competition. Yet here we see a particular focus on individual teachers as a 'front-line workers' (OECD, 2014) who are the best mechanisms to improve educational markets. Vagi (2014)

explained, “by rewarding effective teachers and providing them with opportunities to earn salaries comparable to those in other fields, advocates argue that more highly qualified candidates will be drawn to the profession as less effective teachers are forced to leave” (p. 99). In other words, advocates suggest over time incentives will provoke strong teachers to join and stay and weak teachers to leave the field (Hanushek & Lindseth, 2009). When teachers (i.e. agents) exert their maximum efforts, school districts (i.e. principals) needs are fulfilled and “moral hazards” are reduced (Goldhaber, DeArmond, Player, & Choi, 2008). Because districts can offer attractive compensation, they can retain effective teachers, differentiate effectiveness, and thus increase the average level of teacher effectiveness (Leigh, 2012). Additionally, pay equalizing practices, such as the single salary schedule, reduce the overall performance, attractiveness, and effectiveness of the education labor force and creates high opportunity costs for uniquely skilled people to enter the teaching labor market (Goldhaber, DeArmond, & Deburgomaster, 2011, Podgursky & Springer, 2007b). Lazear (2003) makes clear that even if incentive effects are not strong (i.e., merit pay fails to improve teacher and/or student outcomes), positive selection effects will eventually lead to improved schools. In addition, non-governmental organizations and governments in developing countries suggest incentives can improve exceptionally low teacher salaries, overrepresentation of women, and teacher shortages. Overall, advocates suggest PBC fulfills three key aims: it *alters* the labor force, *maximizes* the labor force’s effort/performance, and thereby *bolsters* educator quality, school quality, and education markets (Lubienski, Scott, & DeBray, 2011).

The role of ideas in GEP reform affords a more nuanced understanding of the structures and institutions that enable the adoption of PBC. It is already well documented that policy actors’ aims often differ from teachers’ perceptions of policies targeted at their improvement.

Acknowledging and outlining the differences between policy actors' and teachers' perceptions on various education reforms helps illuminate global perspectives on high stakes teacher accountability policies. But by tracing these perspectives from the vantage point of the nexus of ideas and institutional change in the Shanghai context, I consider the purchase of the currently fashionable cognitive frames that persuade policy actors to incentivize and monetize teacher performance and worth.

Agency, Expectancy, and Goals

Since the 1960s, scholars have explored the relationships between pay, financial incentives, motivation, and quality of output. While economic perspectives focus on performance, behavioral perspectives focus on behavior. Economists have drawn upon agency theory to explain that extrinsic rewards extract maximum and optimal human performance. Behaviorists, however, have used theoretical tools on expectancy and goal setting to examine the exact behaviors that link incentives and performance. The focus on behavior is key to fully explore how incentives, motivation, performance, and improvement unfold in school settings. Together, these conceptualizations of incentives offer a robust way to consider teachers' perceptions of how bonuses shape their practice.

Extrinsic Incentives Lead to Optimal Performance

The economic principle of agency suggests monetary, explicit incentives, rather than "fuzzy," social extrinsic motivators (e.g., fear of termination, popularity, respect), motivate individuals to expend their maximum effort and behave in ways that lead to optimal performance (Kreps, 1997; Weiner, 1980). Economists further note that the relationship between incentives and human motivation is both irrelevant and uneconomic. Instead, of greater concern is the scope and design of incentive programs (Baker et al., 1988; Earn, 1982). Level (i.e., the total cost of a

package for an employer or the total value of a package to an employee), functional form (i.e., the performance incentives), and composition (i.e., the relative amounts of the discrete components of the package such as cash, retirement benefits, travel reimbursement) are three key issues that labor economists suggest are vital to promising performance-pay schemes. Functional form is particularly important because it bears directly upon the level at which employees value packages (Baker et al., 1998).

Goldhaber and colleagues (2008) framed the rarely used principal-agent relationship in the context of incentives in education and acknowledge the main challenge is school districts (i.e., principals), do not know enough about teachers' (i.e., agents) work to create an incentive scheme that yields maximum payoffs.

> …The central problem for the principal is to structure an incentive scheme that will persuade the agent to act according to the principal's aims so that the principal's expected utility of the payoff is maximized (Dixit, 2002; Laffont and Martimort, 2002). Though rarely used in education, principal-agent theory offers insight into why a school district would choose to offer merit pay or not. Districts, like most employers, face a moral hazard problem: they have less information about the work teachers do than teachers have themselves. Principal-agent theory considers relationships in which one actor (the principal) wants another actor (the agent) to act on his or her behalf. The relationship represents a contracting problem in which the principal must pay for the agent's effort, which produces an outcome that affects the principal's payoff.

Using a principal-agent perspective to model merit pay in school districts, the authors conclude districts are less likely to offer merit pay when the political costs of implementation are high, districts are more likely to offer merit pay when there is more performance information about teachers, and teacher pay will likely be higher in districts with merit pay than in those without merit pay.

Advocates also suggest when districts offer attractive compensation, they can retain effective teachers, differentiate effectiveness, and thus increase the average level of teacher effectiveness (Leigh, 2012). Vagi (2014) explained, "by rewarding effective teachers and

providing them with opportunities to earn salaries comparable to those in other fields, advocates argue that more highly qualified candidates will be drawn to the profession as less effective teachers are forced to leave" (p. 99). Additionally, pay equalizing practices, such as the single salary schedule, reduce the overall performance, attractiveness, and effectiveness of the education labor force and creates high opportunity costs for uniquely skilled people to enter the teaching labor market (Goldhaber et al., 2011; Podgursky & Springer, 2007b). Lazear (2003) asserted that even if incentive effects are not strong (i.e., merit pay fails to improve teacher and/or student outcomes), positive selection effects will eventually lead to improved schools. In addition, non-governmental organizations and governments in developing countries suggest incentives can improve exceptionally low teacher salaries, overrepresentation of women, and teacher shortages (Kobakhidze, 2010). Overall, advocates suggest PBC *alters* the labor force, *maximizes* the labor force's effort/performance, and thereby bolster educator quality, school quality, and education markets (Lubienski et al., 2011).

Perspectives of Pay Matter

Focusing on the individual behavior that links rewards and performance, behavioral perspectives add further depth to economists' reward-performance thesis. Early behaviorists attempted to disentangle the effects of financial rewards upon intrinsic motivation from the effect of financial rewards upon performance outputs (Weiner, 1980). Since the 1960s, when behaviorists first reported financial incentives decreased intrinsic motivation and creativity, no experimental evidence has found merit pay to positively impact intrinsic motivation (Weiner, 1980). This research also uncovered a number of concerns associated with merit pay including narrowing of task focus, disproportionate focus on quick results, aversion to risk-taking, erosion of intrinsic motivation, and the feeling of being controlled by a reward (Kohn in Baker et al.,

1998). Behaviorists have suggested supervisor feedback, perceptions of pay, pay satisfaction, and trust all mediate intrinsic motivation (Lawler, 1983, Luthans & Stajkovic, 1999), and employee participate in compensation design in one way to decrease information asymmetry and increase employees' sense of control (Lawler, 1976).

Expectancy and goal setting are two particular behavioral theories used in recent studies of performance incentives in education. Expectancy theory provides a lens through which one can examine the potency of incentives to lead to improved or desirable behaviors. This framework suggests an individual's perspectives of expectancy (i.e., effort leads to reward), instrumentality (i.e., performance leads to reward), valence (i.e., attractiveness of the expected rewards), are key factors that drive motivation (Heneman, 1998; Vroom, 1964). Recent research finds school leaders' perceptions of expectancy and instrumentality influence where they allocate their effort (e.g., professional development, content, teaching methods), and school leaders' perceptions of valence influence school improvement agendas (Kelley et al., 2000). Also Rice and colleagues (2015) found perceptions of different aspects of fairness (i.e., procedural, substantive, and distributional) shape expectancy, instrumentality, and valence. Goal setting theory suggests factors of fairness and goal attainability also mediate individual motivation and capacity for improved or desirable performance (Bryan & Locke, 1967; Locke & Latham, 2006). In their use of goal setting and expectancy theories in their analysis of a performance-pay program, Rice and colleagues (2015) explained:

> Taken together, studies drawing on these two theories suggest that (a) awards must have high valence, meaning they must be salient and sizable enough to appeal to teachers; (b) awards must address expectancy and instrumentality (i.e., attainability) by demonstrating understandable and credible connections between work, performance, and reward; and (c) goals, measures, and awards must be perceived as fair (p. 30)

CHAPTER 4

RESEARCH DESIGN

This study addressed how merit pay was perceived, functioned, and shaped teaching in Shanghai. In this chapter I describe my constructivist approach to this multilingual, multisite qualitative case study. I also describe in detail the research questions and corresponding data collection and analysis procedures summarized in the below table. I also explain the translation procedures that I employed in my analysis of interview data, and I identify the limitations and strengths of this study.

Table 1. Overview of Research Questions and Methodological Approach

Research Questions	Data Collected	Analytic Methods
1) How do the stated and perceived aims of performance-based compensation (PBC) vary?		
1a) What are city, district, and school policymakers' stated aims of PBC? 1b) What are teachers' perceptions of the aims of PBC? 1c) How do teachers define PBC? 1d) What do teachers perceive to be the desired behaviors or outcomes from PBC? e) How do the stated aims and perceived aims of performance-based compensation vary?	20 teacher interviews, 1 principal interview, 4 policy documents	Inductive and deductive coding, matrix analysis, constant comparative analysis by participant type – policymaker, administrator, novice teacher, veteran teacher
2) How does culture mediate the enactment of PBC?		
2a) What is the role of Confucianism? 2b) What is the role of school culture?	20 teacher interviews, 1 principal interviews, 4 policy documents	Inductive and deductive coding, matrix analysis
3) How does performance-based compensation (PBC) shape instruction?		
3a) How does PBC shape motivation? 3b) What is the relationship between effort and performance?	20 teacher interviews	Inductive and deductive coding, matrix analysis

Research Design Rationale

As mentioned earlier, until recently PBC has been adopted at a local level and for short periods in school districts that grapple with poor educational outcomes. While England and Wales have adopted a national PBC system, it has only been in place since 2013, and it has not profoundly penetrated the system. Also, England and Wales have neither led in comparative international assessments nor are seen as "reference societies" in GEP. These countries are examples of those that are just beginning to situate education policy in larger domestic policy goals related to economic improvement and innovation (Henig, 2013).

Studies of merit pay have largely focused on impact over implementation. These econometric studies generally ignore the policy, political, organizational, sociocultural, and behavioral variables that shape merit pay implementation and impact. In addition, econometric, context-neutral approaches do little to interrogate the fact that teacher accountability policies, such as merit pay, are expected to perform well regardless of place, context, or interpretation (Sahlberg, 2011). Such thin analyses of merit pay pilots have little leverage in advancing our understanding of incentivism in the context of high stakes teacher accountability in GEP.

Scholars who have documented the need for context-sensitivity note that the efficacy of accountability policies can only be understood through multi-focal investigations of impact *and* implementation that explicitly attend to policy actors' perceptions of said policies (Glewwe, 2014; Verger, 2014). Contexts in which high stakes accountability policies such as merit pay are enacted on a state or national level, for a significant period of time, offer a more comprehensive landscape to study policy performance. To understand the capacity of incentivism to positively affect teaching and learning, research must look at subjective variables, such as motivation and self-efficacy. Research is needed that methodologically consider impact *and* implementation,

honors complex context, and examines regions that are embracing teacher accountability policies on a comprehensive level. Shanghai, China's public schools are an ideal case for such research.

Constructionist Approach to Multilingual Research

In all qualitative forms of research, and particularly in interviews, language is the means through which we investigate, interpret, and theorize on phenomena. As Inhetveen (2012) explained, "language is crucial for all methods that aim at tracing the participants' generalized patterns of meaning and experience" (p. 29). I employed a constructionist, ecological approach (Temple, 2002; Temple & Edwards, 2002; Temple & Young, 2004) to this multi-lingual research. I exercised three particular tactics in this approach: 1) partnerships with co-researchers, 2) honoring Confucian norms, and 3) translation for meaning.

Partnerships With Co-Researchers

In cross-cultural and multilingual research, interpreters bear great influence upon the administration and translation of data. Interview content widely differs based upon the discourses and narratives that the interpreters, interviewees, and researchers assemble during the dialogue (Rosenblatt, 2012). In turn, interpreters must be removed from the role of "shadowy figure" (Temple, 2002) and considered legitimate co-researchers in all phases of data collection and analysis. Relatedly, the articulation of the qualifications and background of the interpreters lends to greater credibility and trustworthiness of their "hybrid" roles as "analysts and cultural brokers" (Matteson & Lincoln, 2009; Larkin, Dierckx de Casterlé, & Schotsmans, 2007).

To accurately and carefully interpretation the interviews, I worked with three co-researchers from East China Normal University (ECNU) and used a dialogic translation and interpretation procedures. I established partnerships with three (as opposed to just one) co-researchers for a few reasons. First, working with more partners was the best solution for a

limited data collection period of just four months. Second, translation by committee reduces loss of clarity, increases congruence (i.e., word equivalence), and offers opportunities to compare versions of interpretation from multiple viewpoints that are informed by interpreters with various backgrounds and experiences. (Esposito, 2001). In addition, my co-researchers were interested in educational research, and this project gave them a platform to learn qualitative research methods.

My co-researchers were Lu Keyi, Yang Yang, and Zheng Xue, three, early 20s, female Master's students of teacher education at ECNU. They were students of Professor Jiacheng Li, Professor of Curriculum Studies and Co-Director of the ECNU Institute of School Reform. Professor Li's research focuses on the educational opportunities of internal migrant students, the teaching practices of *banzhuren* (i.e., the head class teachers who impart character and civics education), and parent-teacher partnerships in innovative homework routines. Professor Li has partnerships with over 100 schools in Shanghai and also works with schools in cities in China such as Changshou, Wuxi, Guizhou, Xiamen, and Shenzhen. He has hosted numerous conferences at ECNU that have featured well-known scholars in China and scholars from foreign institutions in the U.S., U.K., Sweden, Australia, The Netherlands, and Finland. Professor Li also regularly collaborates with faculty from Vanderbilt University and University of Michigan. Alongside Professor Li throughout their three-year Master's program, these students collect data in schools, analyze data, and writing academic articles for his various research projects.

In spring 2015, Professor Li introduced me to Keyi, then a first-year student, who joined me during the research design development phase. In fall 2016, Keyi introduced my project to Yang and Xue, two of Professor Li's new students. Since fall 2016, the four of us have worked closely together. Originally from Zhejiang provence, Lu Keyi attended University of Nanjing for her Bachelor's degree in secondary education. Zheng Xue is originally from Sichuan provence

and earned her Bachelor's degree from Sichual Normal University in elementary education, and Yang Yang is from Guangdong provence and earned her Bachelor's degree from Zhejiang University in secondary education. As highly selective universities, these three institutions typically select the top five percent of high school students in China. All three students completed a student teaching practicum as part of their Bachelor degree requirements, passed English proficiency exams with the highest honors, and they matriculated directly into the ECNU Master's program in Pedagogical Studies. The three students come from what is considered lower middle-income families in China. Their parents have educations from universities in China, and they held jobs in government offices and medium-size companies. As is common with university students and faculty, all three students and their families are members of the National Communist Party. My co-researchers mediated interpretation, gave insider knowledge to language culture, were included throughout research design, and they were recognized in this reporting of research. Playing the role of cultural conduits, transcriber, interpreter, translator – these partners and coresearchers pave the way for the rigorous, trustworthy analyses needed in cross-cultural research studies of global education policy.

Honoring Confucian Norms Within Anglo Qualitative Research Traditions

Park and Lunt (2015) explained that the Confucian tradition of saving face is "endemic to the challenge of qualitative research" because participants' desires to produce "socially desirable answers" is inevitable (p. 11). In reflecting upon the clash of Confucian norms with Anglo research traditions in their research on performance-pay with South Korean civil servants, Park and Lunt reported, "prior to the interview, several participants said they had many complaints about incentives, but during the interview their answers were far more guarded…the interview may still have difficulties in obtaining clear perceptions, particularly in organizational contexts

where Confucianism can constrain by producing organizationally desirable responses" (p. 11). Their critical question, "how are such (Confucian) normal balanced are displaced by those associated with western qualitative research at the stage of analysis," (p. 11) guided a great deal of my thinking. One illustrative example that the researchers provided considered how a Western tradition of comprehensive, systematic analysis of qualitative interviews that presents a balanced view of findings and honors all perspectives, even critical ones, was somewhat at odds with Confucian traditions of public displays of optimism and communal goals. The balance and privileging of cultural normal and research paradigms was an issue I considered throughout research design, not only during the analysis phase.

Fontana and Frey (2005) explain that researchers must consider several issues in qualitative research design. Researchers must consider access to and location of the site(s) and participants, participant language and culture, trust, rapport, and collection of key materials related to data (e.g., audio, video, photography, documents, artifacts). In my weekly meetings with my co-researchers, we discussed these issues from a Confucian stand point. School M1 was 20 miles or 80 minutes (by metro and taxi) from my home. I was an outsider in my own neighborhood but particularly foreign to these school surroundings.

Most of the teacher participants were Shanghainese and thus spoke both *Shanghaihua* and Chinese Mandarin, *Putonghua*, as primary languages. Confucian values were deeply embedded in their approach to teaching and learning, sense of professional responsibility, and their way of participating in society (Park & Lunt, 2015). Teachers' preexisting relationships with Professor Li and my co-researchers helped establish a strong foundation of trust with the teachers (Garton & Copland, 2010). Also, Shanghai teachers are accustomed to school-based research. Small and large-group observations of teaching (i.e., lesson study) and action research

are longstanding traditions in this Confucian culture. ECNU-facilitated introductions to principals coupled with my status as an Asian female who formerly taught elementary grades and who speaks limited *Putonghua* created an atmosphere of inside research. However, the variables that created an insider experience (e.g., gender, ethnicity, profession) were present alongside variables that created an outsider experience (e.g., citizenship, ethnicity, language). No one single identity marker superseded another, thus reflecting a more nuanced insider-outsider research among my co-researchers and me (Mercer, 2006).

Translation and Calibration of Meaning

Scholars who reject a positivist approach and advocate for a constructionist lens cite a number of issues in the translation and interpretation of multilingual research. Feminist researchers Temple (2002) and Edwards (1998) offered a social constructionist approach to translation that departs from a quantitative, positivist view that word equivalence requires forward-backward translation. A positivist paradigm sees interpreters as something to "control for" and a "threat to validity" and thus often strives to make interpreters "invisible in the process and product" (Berman & Tyyskä, 2010, p. 179). By viewing research as "language free", this approach presumes congruence (i.e., equivalence) and clarity (i.e., meaning), retains one-to-one correspondence in source and target languages (Larkin et al., 2007). A constructionist approach forces researchers to consider "serious questions about ambiguities and ownership of translated language content; assumptions about community familiarity and cultural similarity between researchers, translators, and participants; negotiation of power and authority in the research process; and the risks faced by translators" (Berman and Tyyskä, 2010, p. 186).

Prioritize the transfer of meaning over equivalency. Matteson and Lincoln (2009) warned that as researchers code and interpret language, their own voices can "drown out those of

the participants" (p. 659). Temple and colleagues (2006) explain cross-cultural research is fundamentally reconstructed primary data and bears great resemblance to the examination of secondary data and thus holds great barriers in terms of reflexive interpretation. Simon (1996, in Temple et al., 2006) comments underscored the contentious nature of word equivalence:

> The solutions to many of the translator's dilemmas are not to be found in dictionaries, but rather in an understanding of the way language is tied to social realities, to literary forms and to changing identities. Translators must constantly make decisions about the cultural meanings which language carries, and evaluate the degree to which the two different worlds they inhibit are 'the same'. These are not technical difficulties, they are not the domain of specialists in obscure or quaint vocabularies. They demand the exercise of a range of intelligences. In fact the process of meaning transfer has less to do with finding the cultural inscription of a term than in reconstructing its value (p.7)

In contrast to English-speaking North America's low-context culture, linguistic meaning is more deeply engrained in Chinese culture, and thus analysis of interview data includes both translation and calibration. Jagosh and Boudreau (2009) explain, "Whereas translation involves finding equivalency between source and target languages, calibration explores whether a word has the same placement or weight in the linguistic field of the target language as the source language" (p. 105). Temple and Young (2004, in Jagosh and Boudreau, 2009) further explained, "the lack of one-to-one relationship between language and meaning does not absolve the researcher from investigating the role of language in cross-cultural research. Instead, it indicates that the boundaries of language are permeable" (p. 174). Nontransferability of language is thus neither a limitation nor a failure but rather a key element to and the essence of a social constructionist approach to cross-cultural research. To calibrate and illuminate areas of nontransferability of language, or language resistance, piloting and pretesting research instruments is helpful (Jagosh & Boudreau, 2009).

In addition to Temple and Young's (2004) methods of researcher-translator (i.e., researcher acts as translator) and researcher plus translator (i.e., researcher outsources translators

to independent, external personnel), Fersch (2013) proposes we consider that oftentimes researcher-translator plus interpreter. Oftentimes researchers who are translators must interpret in a language in which they are fluent but is not their native tongue. As such, word equivalence, clarity, courtesy, and calibration become a challenge.

Scholars have suggested three analysis activities to ensure translators "transfer the meaning of the words rather than re-writing them" (Al-Amer, Ramjan, Glew, Darwish, & Salamonson, 2014, p. 1159) and to increase the overall trustworthiness of data interpretation: debriefing sessions with co-researchers, independent verification of the accuracy of translation, and construction of independent categories to address word equivalence (Temple et al., 2006). Four dimensions of translation – cohesion, congruence, clarity, and courtesy (Larkin et al., 2007) – guided our data collection and interpretation procedures. Larkin and colleagues explain cohesion reflects the balance of interpreter-researcher, interpreter-participant, and researcher-participant relationships, and as well as the interpreters' cultural and linguistic knowledge and the researcher's methodological expertise; congruence denotes the linguistic equivalence (i.e. word, syntax, grammar, concept) between the target and translated languages; clarity attends to the quality of detail and meaning across the languages; and courtesy represents the balance between formal and informal data and appropriate application of cultural decorum.

Case Study

To understand the implementation and impact of incentivism on teachers, I employed purposeful stratified sampling in a qualitative, multisite case study (Yin, 2003). Multisite case studies are a useful way to expand our incomplete understanding of complex, interrelated phenomena such as incentivism and teaching quality (Yin, 2003). Investigating the phenomenon

of incentivism in multiple cases was important because the model has not been fully conceptualized, as others and I have argued (Malen et al., 2015; Verger, 2014).

I investigated incentivism in the form of merit pay (i.e., performance pay, pay-for-performance), or financial bonuses that led to greater compensation and were determined by various performance metrics, including but not limited to student performance on standardized tests of core subjects. This form of enticements is the most pervasive in school autonomy policies today and is advocated for by the OECD and leading world economies (OECD 2009; United States Department of Education, 2015). My methods included semi-structured interviews and document analysis. In this multisite case study, I interviewed 20 teachers and the principal at Schools M1, a high performing elementary schools in Minhang District in Shanghai. I also reviewed school, city, and national documents on merit pay. In this section I describe the data collection and analysis that occurred in detail. I refer to the faculty and students from East China Normal University (ECNU)who supported the collection, translation, and interpretation of data throughout the duration of this research as my co-researchers. The participant breakdown, interview protocol, and a list of key terms are all located in the appendix.

Purposeful Stratified Sampling

In school year 2014-15, Shanghai had 764 public schools with 798, 686 students and 52,321 teachers. The budget for expenditure on public schools is 10,961,715 USD (76732 million RMB), which represented 12.4 percent of the total budget (Statistical Yearbook, 2016). The public schools of Shanghai are divided into 17 districts, which are located in the urban core, inner suburbs, and the outer suburbs. Nine districts comprise the urban, original core of Shanghai: Huangpu, Luwan, Xuhui, Changning, Jing'an, Putuo, Zhabei, Hongkkou, and Yangpu. There are nine inner and outer suburban districts, of which the fastest growing is Minhang, and

Chongming island is a district accessible by bridge. The 2014-2015 school year rankings were published on November 11, 2015. Each district designated first, second, third, and fourth place to several elementary, middle, and high schools. Compulsory schools with grades 1-9 were placed in a separate category. A map of Shanghai and a list of the 2014-15 Minhang District school ranking are located in the Appendix.

In July and August of 2016, I employed a purposeful and stratified sampling technique (Creswell, 2003) to select schools and participants for a multisite case study. In order to investigate merit pay enactment and impact, I targeted schools that were high performing and had robust merit pay programs. These criteria offered opportunities to test two intersecting hypotheses: 1) merit pay will increase an agent's motivation to improve, thereby leading to optimal behavior on the part of the agent and optimal output (Springer, 2009; Weiner, 1980); and 2) high performing schools, such as those in Shanghai, are models for education reforms and practices that lead to high quality teaching and learning (Burningham, 2014; Morrison, 2014; Schleicher, 2016; Verger, 2014). School performance was determined through publicly available school rankings on the Shanghai's district's websites. The 2014-2015 school year rankings were published on November 11, 2015. School performance was determined through publicly available school rankings on the Shanghai's district's websites.

My co-researchers suggested that given the sensitivity of the policy, it was best to target highly ranked elementary (grades 1-5) and compulsory (grades 1-9) schools in Minhang District, as the ECNU faculty members had preexisting relationships with many schools in this area. In summer of 2016, via e-mail, my co-researchers and I contacted the principals of 13 elementary schools in Minhang ranked in first and second place. We sent an introduction e-mail and a short survey of interest that addressed the selection criteria of a high performing school and a robust

merit pay program. The survey of interest determined whether merit pay was fully implemented for at least five years, comprised a substantial amount of teacher salary, and whether student performance on standardized tests was a merit pay measure (see Appendix).

Three schools responded with interest in participation. When I met informally with the principals of these schools, I answered their questions about the study and conveyed my commitment to anonymity, confidentiality, and reciprocity in the form of professional development. I also asked additional questions to ascertain whether the school had a balance of veteran and novice teachers in terms of overall experience and experience in that particular school. In addition to the criteria of high performance and robust merit pay, teachers with varied years of experience in the school and in the profession was important for obtaining varied perspectives on purposes of merit pay and the policy's impact on teaching. The Principal of School M1 offered the greatest access to teachers and was open to a long-term research project.

School M1. In school year 2014-15, there were 65 schools in Minhang. M1 ranked in first place in each district. Upon the merger of two nearby elementary schools in the district, M1 was reorganized and renamed in 2012. In 2014-15, M1 had 1,200 students and 80 teachers. The principals and vice principals reported that the schools employed teachers with a range of one to 30 years of experience as educators and experience in the school, and the average years of teaching experience among the teachers was ten years.

M1 offered seven academic subjects – art, Chinese, English, mathematics, music, physical education, and science, and it held the the designation of *experimental schools*. *Experimental schools* in Shanghai typically have a record of strong student outcomes and actively participate in research partnerships with universities and serve as incubation sites for organizational, pedagogical, and curricular reforms. M1 received district and city accolades for

their Young Pioneers (i.e., red flag brigade) program, teacher professional development programs, civics education, and teacher moral education training, and it is a teacher demonstration school. M1 followed the 2009 city mandate to implement a 70-30 percent base pay-merit pay program in their schools. The teachers received a monthly salary that included both base and merit pay.

Collection of Interview Data

In fall of 2016, using an interview guide, my co-researchers and I conducted semi-structured, one-to-one interviews (Bryman, 2004; Flick, 1998) with 20 teachers at M1 and with the principal.

Protocol. In January of 2016, I created an interview guide based on Patton's (1990) framework, using informal, open-ended, and more formulated questions. Previous studies and theory on incentivism and teacher development informed this protocol. In spring 2016, I piloted the interview guide in Chinese Mandarin and in English with my co-researchers. After we refined the major and probing questions, we met with a focus group of five elementary school teachers and tested the translation of keywords in Chinese Mandarin to check for for consistency in meaning. The Appendix contains a list of Chinese Mandarin and English key terms that were used frequently during the the interviews.

The four sets of questions in the interview guide elicited participant perspectives on the aims of merit pay, implementation of merit pay, impact of merit pay on motivation to improve teaching, and impact of merit pay on relationships. The cognitive method of probing was employed in order to help verbal material that would be useful for congruence and clarity in the translation process (Matteson & Lincoln, 2009). Each interview question thus included at least

four scripted probes for the co-researchers to choose from, and we used spontaneous probes throughout the interviews, also.

Participants. The M1 Principal requested to facilitate the organization of teacher interviews according to my selection criteria. This request was common in school-based research in Shanghai because of the large faculty sizes and block scheduling approach in the schools. During the information gathering interviews, the principals offered to set up the interviews according to my criteria. I interviewed 20 teachers, approximately 25 percent of the faculty, in each school. The participant sample reflected diversity by grade level, subject (Chinese, English, mathematics, science, art, physical education, music), years of teaching experience, years in the school, age, and amount of merit pay received in the 2015-16 school year. Because approximately 90 percent of elementary school teachers in Shanghai are female, it was understood that the sample would include relatively few males. An inventory of the participants is located in the Appendix.

Structure. Interviews lasted between 20 to 60 minutes each. Interviews with principals took place during the school day in each principal's respective office. Interviews with teachers took place during teachers' preparation periods (with principal permission) and after school hours. One co-researcher and I conducted and audio-recorded (with consent) all interviews in person in a large meeting room at each school in Chinese Mandarin. I interviewed each teacher once and the Executive Principal twice. The first interview with the principal focused on understanding their perceptions of the aims of merit pay, policy enactment, and the ways in which incentivism shapes teachers' motivation to improve instruction. The second interview with the principal probed for clarifications on merit pay policy as written and as implemented and the capacity of incentives to shape instructional improvement. Teachers were asked about their

perceptions of financial bonuses, the impact of bonuses on instruction and collegiality, and their personal experiences with incentives. I contacted teachers via text message or via e-mail with follow-up questions. In Confucian cultures such as Shanghai, administrators and teachers often meet with researchers during working hours or on the weekends, and all participants receive a small gift and refreshments for their participation (Park & Lunt, 2015). Upon the recommendation of my co-researchers, I gave the principals and all teachers participants small gifts worth approximately five dollars per item.

Collection of Document Data

In spring of 2016, I gathered publicly available documents on merit pay policy in Shanghai. Blog posts and web-based news articles offered background on the climate around PBC and accountability reforms in Shanghai. I also collected research-based accounts of merit pay in Shanghai schools. These studies were published in non-refereed journals and written as Master's theses and academic journal articles. In addition, I collected publicly available city and national policy documents on merit pay, and the M1 principal gave me school merit pay policies. Most documents were published in Chinese Mandarin. I used one document review template for the popular and university-based research documents. In this template I recorded a summary of content, and, as applicable, performance incentives aims cited, findings, and relevant citations.

Data Analysis Procedures

Goal-setting, expectancy, and agency theory guided the initial set of codes used to analyze the interview data. Such a multilayered view of individual behavior was necessary to fully capture both policy aims and participants' interpretations of financial bonuses. All interview data were transcribed in Chinese Mandarin, translated to English, and then and analyzed to address the central questions motivating the study. The analysis of the interview and

policy document data occurred in four steps: translation, coding and memoing, check for disconfirming evidence, and review of interpretation.

Translation. My co-researchers transcribed the interviews and the school, city, and national merit pay policy documents. Whenever possible, the co-researcher transcribed the audio recordings of the interviews in which she was present. To ensure consistency in the interpretation of specialized terms, the translators used the list of key terms in Chinese Mandarin and in English. In each transcript, we searched for the key terms in Chinese and checked for accurate English translation of the term in the context in which it was mentioned. We consulted one another and corrected inconsistencies throughout this process. I met with ECNU faculty in the Department of Governance and Management to review the policy translation and discuss meanings and processes surrounding key implementation terms.

Coding and memoing. I coded the interview data inductively and deductively. I coded the entire dataset once inductively; then used a constant-comparative analytic approach (Glaser & Strauss, 1967) to modify, delete, replace, or expand upon these codes based upon sentiments evoked by the participants; and then I coded the entire dataset a second time. Throughout this process I wrote and refined memos on emerging themes. Each memo contained excerpts of interview data that supported and refuted the emerging theme, a list of clarification questions, and relevant school context. I also maintained a list terms in Chinese Mandarin that were often used and required further refinement in English translation and a list of questions for investigation in future research.

After the first round of coding I wrote a memo on emerging themes for each research question. Each memo contained excerpts of interview data that supported and refuted the emerging theme. I also included a list of clarification questions on these themes and on school

context that were relevant to the themes. I revised the memos after the second round of coding. I created a separate memo of a list of terms in Chinese Mandarin that were often used and required further refinement in English translation. I also created two flow charts – one to describe the policy processes outlined in the policy documents and by administrators, and one to that highlighted policy implementation as characterized by teachers.

Check for disconfirming evidence. I selected a random sample of one page of text in five (25 percent) out of the 20 teacher interviews in order to check for disconfirming evidence (Miles & Huberman, 1994). My co-researchers and I coded this random sample. There was 90 percent agreement in assigning a given code. We also performed a syntax analysis on this random sample to analyze the descriptive words that teachers most often used when they described their perspectives on merit pay impact and implementation. We used this list of high frequency words to further refine the memos of findings and build depth to the findings. The constant comparison, keywords in context, and domain analysis procedures improved the reliability and trustworthiness of the data (Leech & Onwuegbuzie, 2007).

Strengthen interpretation of findings. I shared my memos with an ECNU professor of teacher education who works with over 100 schools in Shanghai. I also held data interpretation meetings with the principal and one teacher in each school. During these meetings, I shared condensed versions of the memos and gave a short presentation of findings. The participants added rich context on cultural and micro institutional variables that shaped merit pay perceptions and policy implementation.

Strengthen interpretation of findings. I shared my memos with two professors of teacher education at ECNU who work with teachers and principals on curriculum and instruction in more than 100 schools in Shanghai. I also shared my memos with the principals and one teacher in

each school. During these data interpretation meetings, I explained findings and offered several examples of interview data that exemplified each finding. The principals helped refine the policy implementation flow chart and added rich context on micro institutional variables contributed to differences between policy design and implementation. The teachers added rich context to explain the cultural underpinnings of findings related to impact on teacher practice.

Strengths and Limitations of Research Design

Readers benefit deeply from qualitative case studies because this approach is most harmonious with our natural way of learning things – through experience. To generate in-depth understandings of how incentives shape teachers, students, and schools, qualitative case study offers an accessible documentation of experience. Still, this was a study of multiple schools in just one region. These findings are neither generalizable nor capture all views within the site itself. Relatedly, I was not in these schools for extended periods of time. Interview data were collected in a relatively short time period due to the extensive time required for transcription and translation. Ethnographic methods that included observations and debriefs likely would have lent to a more thorough representation of the context, climate, and culture of the schools and participants.

There are also limitations to the interview method both generally and as applied to this case study. Aiming to capture participants' perceptions of overall patterns of interaction and themes rather than specific, isolated incidents, I asked open-ended questions and then probed for examples. I also used member checks frequently. However, because of my limited Chinese Mandarin proficiency, it was challenging to clarify and probe for meaning when participants' self-reports were conflicting, vague, or even particularly intriguing. I chose to honor the flow of communication and mimic as natural of a conversation as possible rather engage in translation

throughout the interviews. Instead, my co-researchers and I isolated a few probing questions from the interview guide and from our own thoughts and spent the last five minutes or so on clarifying or expanding participants' responses.

Finally, and perhaps most importantly, by using the English translations of interview transcripts rather than the Chinese Mandarin transcriptions, it is possible I inadvertently drew upon my tacit knowledge of school improvement, teaching, and learning in my interpretation of participants' views. It is also possible that through translation, the nuanced and delicate details around policy implementation were omitted or ignored, and participants' unique feelings and interpretations of bonuses were misunderstood or underemphasized. Testing for non-confirming evidence and data interpretation meeting with teachers helped mitigate these challenges.

Despite these limitations, this study has several important strengths. In comparison to the total number of schools in Shanghai, the number of schools and participants studied is small. However, in comparison to the literature on PBC, this study offers a much needed qualitative exploration of questions that are understudied in a site that is crucial to global education reform. The careful and thorough analysis procedures allowed my mental "organizers" to guide the study but encouraged concepts and concerns to surface as they become acclimated with the case.

Researcher Role

Shanghai has been my home since 2009. I know the streets, food, and the ways of the people. I know what is trendy, traditional, and what generally matters to people. I speak the language well enough to convey myself and to understand others. All of this, plus my Indian phenotype, make me an insider. But Shanghai has always made clear to me, I am a *laowai*, a foreigner, in their Shanghainese and Chinese cultures, and this comes with advantages and disadvantages. In the context of this study, participants likely saw me as a fellow educator who

perhaps had some capacity to express their voice and represent them in ways they could not achieve within the confines of their culture and profession. Yet both they and I knew that I would never fully understand what it meant to be a teacher in Shanghai, and thus I would never fully understand their experiences. As with all research involving human subjects, complex phenomena, and especially in research with multiple languages, I took great care to convey, in a culturally appropriate fashion, my longstanding commitment to understand the everyday and usual of Shanghai, and acknowledge that I was an appreciative guest in their culture.

CHAPTER 5

FINDINGS

In this chapter I report on findings from my investigations of several research questions. I begin with a discussion of findings on policymakers' stated and teachers' perceived aims of performance-based compensation. I then discuss how teachers and administrators enacted PBC in the context of Confucian school culture. Last, I report on findings from my analysis of teacher perceptions on how bonuses shaped their motivation to improve their performance.

I find that a culture of deference, stability, and individual responsibility made it possible for what is essentially a salary supplement system to thrive. City and district policymakers and the principal of School M1 suggested bonuses were intended to incentivize teaching quality. The M1 teacher participants viewed financial bonuses as a mechanism to incentivize increased effort that could differentiate strong from weak teachers but not necessarily yield improved teaching quality. Teachers perceived low valence and instrumentality, high expectancy, and they had mixed views on fairness. In distinguishing effort from performance, teachers suggested bonuses had no impact on their instructional performance.

I argue that rewarding teachers for (more of the same) effort is unlikely to positively shape quality. By employing a vague definition of quality, policymakers, administrators, and teachers remain unclear as to exactly what ends incentives are levied. This vague notion of quality is, however, essential for a standardization approach in global education reforms. Also, linking incentives with positive public recognition and self-efficacy may help lead to the improvements in teaching quality desired by policymakers.

Divergent Perspectives on the Aims of Incentives

Idea Penetration From Nation to Provence to Schools

China's comprehensive domestic policy is built upon the goals of world economic leadership and technological innovation. The national government designated Shanghai as a site for economic and educational innovation. When Shanghai emerged as a free trade zone, its educational system followed with reforms intended to support technological advancements, elevate the quality of teaching and learning, and support the country's aims of world economic leadership. Shanghai's 2001 *First Class City, First Class Education* policy aligned broader goals of leading in the global knowledge economy through teaching reform initiatives (Lai & Lo, 2007; OECD, 2012a; Tucker, 2011). New teachers are mentored by veteran teachers and participate in team-based professional development for at least 50 hours per year, and school leaders have an average of 15 years of teaching experience (OECD, 2012a; OECD, 2012b).

Assess, reward, and differentiate teacher performance. Merit pay was also a key strategy in Shanghai's teaching reform approach. In 2009, the Shanghai Municipal Education Commission implemented a city-wide performance appraisal, which included provisions for performance-pay. Policymakers suggested performance assessment, and pay linked to that performance, would motivate teachers to work harder and develop their professional skills. Performance assessment and pay was directly aimed at improving teachers and schools:

> The aim of the new performance wages allocation system is to guarantee teachers get good treatment and improve the quality of teachers. It also has significant effect on promoting balanced development of compulsory education in Shanghai.

They intended to offer performance pay to *banzhuren*, teacher leaders, and those teachers who had "great achievement." (A new performance wages allocation system of compulsory education school was built in Shanghai, 2016)

City policymakers recommended performance assessment be differentiated by teacher title; assess moral education, teaching, research, and professional development; and reflect

teacher ability and achievements. Policymakers also encouraged principals to offer greater incentives to *banzhuren*, in this case the head teacher for each grade level. In addition, the city called upon school districts to: 1) learn about performance-based evaluation and compensation options from publications posted on the central government education ministry website; 2) create a scientific, detailed plan on performance assessment; 3) design the performance assessment procedures carefully; 4) employ multiple methods of performance assessment; 5) perform performance assessments in accordance with the chosen procedure; and 6) use performance assessment results appropriately (Shanghai Municipal Education Commission' opinions on completing compulsory education school teacher performance assessment, 2009). A combination of qualitative and quantitative assessments, with input from students and families, at regular intervals throughout the year, was meant to "produce more forms to record the results of performance and build [the] personal performance development file for teachers." The results of these performance assessments were to be used for performance pay, teacher promotion, and teacher training.

Performance assessment was intended to be guided by a self-designed teacher "work plan," a peer assessment, and an evaluation from the assessment group, which included the Principal, Chinese Communist Party Secretary, school teacher union chairperson, teachers' congress representative, and a handful of teacher leaders:

> At the beginning of each school year, every teacher should make a plan of their work base on the demands of their position, and give it to school performance assessment group. At the end of school year, teacher should write a brief summary of their work in this year and do a self-assessment. Also he (or she) will get a peer-assessment and the assessment made by the school management. The school performance assessment group will synthesize all the assessments and make the final assessment of the teacher. The result of performance assessment has four levels: excellent, pass, barely pass and fail. Only 10-15% of teachers can reach the excellent level. The result of the performance assessment should be made available to teachers. If the teacher has questions about the result, he (or she) can appeal the result to school performance assessment group and

> education authorities. (Shanghai Municipal Education Commission' opinions on completing compulsory education school teacher performance assessment, 2009)

Each district was to supervise performance assessment and mandated the school teacher union approve the school performance assessment policy under this framework.

District Guidelines. The new performance appraisal applied to all school employees who began work by January 1, 2009. For compulsory education school staff salaries, the districts awarded funding to schools in three areas. Job subsidies, or base salary, referred to the five percent of monthly salary that teachers received according to their job title and years of experience. Workload subsidies comprised seventy percent of monthly salaries and was based upon teachers working hours, which was determined by the number of lessons a teacher taught per week. Performance rewards comprised twenty-five percent of monthly salaries and was intended to "reward school staff's great achievement in different areas,…for overtime,…excellent (student) assessments, *banzhuren* (i.e., head teacher), and teaching and research." Both district mandated school performance pay plans include standards for performance pay as well (Scheme of the Implementation of Performance Wages in Compulsory Education Schools in Minhang District, 2008).

Incentivize teacher participation in activities. Performance-based teacher evaluation encompassed three categories of assessment – professional and ethical conduct (e.g., attendance, dress code), teaching, and student affairs. Merit pay was associated with teacher progress in teaching and student affairs. The school management teams (i.e., executive principal, vice-principal, Secretary of Communist Party, teacher labor union representative) developed merit pay policy that would reward teachers for their participation in student learning and teaching activities. Teacher participation in student development activities was particularly important because the focus on "rich campus life" was a new, critical aspect of the school mission.

According to the M1 Executive Principal, parents in the school community were particularly concerned about their children's happiness, and thus teacher participation in student affairs underscored positive school culture:

> Generally, we feel that the teachers in our school have a passionate and positive attitude towards work. The overall work environment is truly excellent. Many teachers from other schools would say, how do you manage to keep doing this, because we cannot go along with it. We think things should be like this. For example, when we organize a student activity such as "four seasons on campus", every one of us will fully participate. Many teachers will provide ideas regarding the carry-out of the activity. This is our culture. We feel that it should be like this. (School M1 Executive Principal, personal communication, September 2, 2016)

M1 offered merit pay to individual and groups of teachers for their participation in student and teacher competitions (see Key Terms); teacher training sessions, lectures and seminars; and in school-based research. A financial incentive and performance-based evaluation points were assigned to each activity, and in the case of competitions, to ranking levels (e.g., first place, second place). An incentive and points ceiling for each activity category was instituted in order to encourage multifaceted teacher development:

> Each category has a maximum score, and the overall maximum score is 20. You can do more, but it will not go over 20. It took us a long time to design this system, because we want the teachers to have balanced development rather than one-sided development, and the maximum score limit encourages them to seek development in multiple aspects. (School M1 Executive Principal, personal communication, September 2, 2016)

Merit pay was also offered to teachers of tested subjects (Chinese, English, and Math) and tested grades (third through fifth) for student performance on benchmark assessments. It was unclear if or how teachers of non-tested subjects (music, art, science, physical education) and grades (first and second) received student performance scores through principal classroom evaluation and parent survey.

The Executive Principal explained performance appraisal and pay helped sort teachers by performance and provoke more effort from underperforming teachers:

> Yes, it needs to be disciplined and democratic, but my review result should never be "everybody is equally good". We need to provide incentives through such evaluation, making teachers who performed well feel recognized, while also making the teachers who performed poorly realize that the school is aware of that. If they see the gap, they will make more efforts next semester.

From her perspective, M1's goal was clear: to identify and reward superior performance through standardized and performance-based measures: "Our principle is those who perform better get more pay. We cannot let the teachers feel that the more they do, the more mistakes they will make, or that they will get the same amount of money whether they work extra or not. They get more when they perform better."

Teacher Perspectives on the Aims of Financial Bonuses

Most of the M1 teachers reported that they were unclear on the intended aims of performance pay but suspected bonuses were intended to positively shape teacher enthusiasm, professional satisfaction, development, and effort. A third-grade Chinese teacher suggested merit pay had the capacity to improve teachers' enthusiasm and served as a guide for improvement. A fourth-grade English teacher elaborated that bonuses helped instill "professional happiness" in all teachers and encouraged teachers to work harder. In terms of specific types of work, she suggested merit pay was intended to encourage teachers to participate in research projects. A male third-grade information technology teacher echoed this sentiment. He added that before 2009, "all teachers got the same pay no matter what they did. It was easy to become lazy."

Teachers elaborated that bonuses were directly tied to evidence of additional work time, yet perspectives differed on the significance of the bonuses. A second-grade *banzhuren* teacher likened merit pay to the work of a farmer: "It feels like a farmer's harvest. However much you cultivate is how much you gain." She emphasized, "merit pay encourages everyone to do their best to accomplish more work." She reported her base salary as about 643 dollars per month and

between 8,500 to 10,000 dollars per years. Several teachers added that bonuses associated with participation in activities provided marginal salary supplements. A first-grade math teacher explained, “There is no ‘high salary’ and ‘low salary.’ At most, they will give you a little extra money to reward you. Now, we both (highly active and less active) teachers get more salary and merit pay. We get more money than before, but there is little difference among teachers.” The physical education department head for third-grade offered a different perspective. She reported large gaps in performance rewards: “Some people get a lot of merit pay, but some people only get a little.” She explained that bonuses ensured “talented and active teachers are valued” because previously those who “did more things” were not compensated for their extra effort. This teacher reported her salary as between $18,000 to $20,000 per year. When probed for salary figures, every teacher said they were unsure of the amount of monthly merit pay they received because this figure was not designated clearly in their pay stubs. Most teachers reported a range of between 400 to 700 dollars total in bonuses per month, with a few teachers declining to comment.

Whether or not bonuses were significant, the majority of participants suggested less active, “lazy,” or “low performing” teachers had the most to gain from financial bonuses. A fifth-grade science teacher commented, “For good teachers, merit pay makes no difference. But for the teachers who are a little poor, there might be some impact.” A first- and second-grade art teacher explained that active, “good” teachers already engaged in activities to support students and teaching, and merit pay was “just a token of affirmation.” Since 15 out of 20 of the teacher participants were teacher leaders (e.g., subject head, grade-level head, *banzhuren*), they represented the highest performing teachers in M1. All of these teacher leaders echoed the

sentiment that merit pay was likely intended to inspire more enthusiasm and participation in activities from the lower performing teachers.

Incentives in Confucian School Culture: Disguising Salary Supplements as Incentives

As with most other policies, School M1 had autonomy to create their own compensation policy. Like most other schools, M1 created a policy wherein monthly salaries were comprised of 70 percent base salary and 30 percent performance-based salary. The school management team – which included the principal, vice principal, Communist Party Secretary, and Chairperson of the Labor Union – authored the policy. It went through several iterations, and staff were given multiple opportunities to offer suggestions for revisions. Like all other school policy, PBC earned approval from 70 percent of faculty.

Structure and Performance Measures: A Range of Rewards and Multiple Metrics

Teacher compensation was comprised of basic salary and performance-based compensation. Basic salary accounted for 70 percent of total pay. PBC included three parts: school-based individual performance bonuses (20%), school-based team performance bonuses (5%), and a district bonus for individual teachers (5%). Basic salary was distributed via direct deposit into teacher bank accounts on the last weekday of each month, individual and team rewards were distributed at the end of fall and spring semesters, and the district award was distributed once per year. The deposit confirmation text message that teachers received differentiated each type of compensation. Several teachers reported they messages did not make clear how much merit pay they had received.

The criteria for the district bonus was not transparent or made available to teachers. Teachers reported they did not know how districts chose teachers of excellence or whom the district chose in a given year. Basic salary was based upon number of working hours. Working

hours refers to the number of lessons a teacher has taught per month. The participants in this study taught between 11 and 17 lessons per week, between two and four lessons per day, and one lesson lasted 90 minutes. Basic salary was based upon professional title. Teachers had a coefficient of one, teacher leaders of subjects and grade-levels had a coefficient of 1.2, and *banzhuren* had a coefficient of 1.3. Basic salary remained quite consistent, as teachers normally taught the same number of lessons per week. School holidays, special school events, student competitions, and end-of-term testing accounted for variance in basic salary.

The individual and team-based PBC allocation and structure was far more detailed. Individual teachers earned merit pay for three different types of activities, and each activity was associated with different awards. Professional development activities included attendance in seminars, lesson study, peer observation, and curriculum development meetings. Teachers also earned bonuses for participating in action research projects. In these projects, teams of teachers, who were typically unified by subject, investigated pedagogical approaches through book study, engagement with university-based researchers, and lesson study. Lesson study refers to process wherein groups of teachers observe a peer teaching a lesson, debrief on the character of teaching and learning in the observed lesson, and then the teacher receives detailed feedback which is implemented in her future instruction.

Teachers of Chinese, English, and mathematics in grades three through five also earned money for student performance in standardized assessments. These teachers and those of non-tested subjects and non-tested grades also had opportunities to earn bonus pay for student competitions. Teachers took students to district, provence, and state competitions. Different bonus amounts were associated with first-, second-, and third-place rankings in these competitions. Physical education and art teachers in upper primary grades were heavily engaged

in student competitions, though teachers of all subjects reported they participated in student competitions. Teachers also participated in their own teaching competitions. In these events, teachers gave mock teaching demonstrations in order to compete against their district and provence counterparts for demonstration of high quality instruction and student engagement.

Authority on awards. When asked about who was involved in merit pay determinations and the processes through which merit pay decision were made, most teachers spoke about performance appraisal. They identified the individuals and involved in and potential outcomes of performance appraisal. Teacher leaders in each subject and grade-level and the school management determined merit pay allocations. Teacher performance-based appraisal was carried out by a team of ten people, which included the school management, leaders in the teaching research group, and various teacher leaders across campus. Each teacher's subject- and grade-level teacher leader also weighed in on the evaluation. All teachers received performance-based rankings of A, B, C, or D. Typically, teachers who received an A grade also earned the most merit pay. The M1 Principal explained that at least 80 percent of teachers received As, 15 percent received Bs, and no more than 5 percent received Cs. The M1 Principal had final say in all evaluation determinations and bonus allocations.

Allocations: Bonuses As Supplemental Pay

Overall, teachers were ambivalent toward the idea of merit pay as an incentive. One Chinese teachers described merit pay as "a salary supplement," and a music teacher described merit pay as "the same as our salary before but just called something else." A veteran teacher further suggested, "basically they reorganized they salary into something we can be sure about and something we cannot be sure about." Teachers suggested merit pay was not much of an incentive and not particularly attractive in size.

Teachers reported their merit pay as between 285 to 1,800 dollars per semester and their base salary as between 430 to 930 dollars per month. The median merit pay reported was 430 dollars, and the median base salary reported was 715 dollars. Of the eight teachers who had six or less years of experience, six teachers reported their merit pay as approximately one-third of their total monthly salary. All teachers reported their salaries did not fluctuate by more than 140 dollars per month, or about 10 percent of their total salary. All of the twelve participants who worked as teachers in Shanghai before 2009 reported their salaries increased by no more than three percent since the onset of PBC. One teacher explained that in 2008 her total salary was 14,290 dollars, and in 2015 her salary was 18,570 dollars.

One teacher reported that she competed in a talent show in the district, and the organizers gave her 71 dollars for her participation, though this had nothing to do with the school- or district-based awards. This same teacher also reported that she received between 1,142 and 1,285 dollars last school year at the end of the first semester and could not remember how much she received at the end of second semester. One physical education teacher reported she received between 1,715 and 1,860 at the end of each semester in school year 2015-16, which was one of the highest reported merit pay. She echoed other participants' sentiments that there was little difference among teachers and "no such thing as high salary or low salary," but everyone had "more money than before." Several teachers noted they earned the most merit pay from leading student activities and helping their students win competitions. Other teachers explained that the additional salary provided by merit pay neither functioned as an incentive nor as an adequate reimbursement for additional effort and leadership. A mathematics teacher explained:

> For me there is not much difference. You do your job and the school assesses it. The fundamental work is teaching classes, grading homework, and instructing the students. I did this before and I do this now. The other activities have limited scores. No matter how much you accomplish, the highest score you can get is 20. Even if you and me get

> different scores, the money we get is not that different. Just a few dozen or hundred yuan. Like me, I am the leader of the teaching and research group, I will just get a few hundred yuan more than the teachers in my group. So we don't care much about it. We will still do our job that we had to do as a teacher.

Policy Enactment In A Culture of Respect, Stability, and Community

Teachers explained merit pay had neither positive nor negative impact on their teamwork. When asked about how financial bonuses shaped teamwork, a third-grade teacher explained, "In general, we teachers at the grass-roots level accept the standard." Participants reported that they generally did not question the management team regarding performance appraisal or performance bonuses. While teachers offered opinions on merit pay allocation and processes, they were clear "pay doesn't mean much to teachers," and a passion for student and school success trumped any shortcomings of the policy.

This teacher went on to explain that teachers did not discuss pay because their attention was on their classes and students. A music teacher explained that the policy "simply reinforces what we are already doing." This teacher elaborated, "we don't do things secretly. We work together on activities. We always do things together. It's good for everyone. Instead of doing things alone, we work together and then make progress together." One *banzhuren* reflected on the link between teacher relationships and financial bonuses and suggested that organizational structure and norms were negated any potential conflict posed by merit pay:

> Principals structure teams with novice and veteran teachers. Experienced teachers can help the novice teachers, and they can also be encouraged by the ambition of the novice teachers. They think about the personalities of the teachers. They give similar merit pay to everyone, so no one thinks about it too much. The balance is important. People are focused on the teaching and togetherness.

Other teachers echoed these points of view and explained that teachers served to help one another, especially during times of struggle.

Effort Vs. Performance: When Incentives Take Teachers Out of the Classroom

An economic view of bonuses draws a linear relationship between financial rewards and teacher performance, while a behavioral view of bonuses suggests teacher perceptions of valence, expectancy, instrumentality, and fairness mediate the efficacy of bonuses. I assessed each teachers' perceptions of awards against these variables. In most cases, the interview guide elicited responses from teachers that spoke directly to whether they perceived high or low valence, expectancy, instrumentality, and fairness. When it was unclear, I probed for their point of view, and in a few instances, I asked directly (e.g., "Does your teaching performance lead to rewards?" Responses were coded as high and low for each variable according to the rubric outlined in the figure below.

Figure 2. *Teacher Perspectives Incentives Related to Dimensions of Expectancy and Goals*

	Valence	Expectancy	Instrumentality	Fairness
High	The reward is salient, sizeable, and appealing.	There is a strong likelihood that effort leads to reward.	There is a strong likelihood that performance leads to reward.	The goals for which rewards are enacted are appropriate. The measures used to determine bonuses are fair.
Low	The bonus is insignificant, slight, and unattractive.	It is unclear or unlikely that effort leads to reward.	It is unclear or unlikely that performance leads to reward.	The goals for which rewards are enacted are amiss. The measures used to determine bonuses are unclear or unfair.

The majority of participants perceived valence and instrumentality as low and perceived expectancy as high. Perspectives on fairness were mixed – some participants found the goals and measures associated with bonuses as fair, some found these unfair, and some teachers conveyed neutrality. Teachers saw their practice, or their teaching performance, as inclusive of their instruction and participation in individual and team-based activities. No matter what perspective they conveyed on fairness, all teachers suggested bonuses had little impact on their instruction but were directly linked to their participation in activities.

Teaching Is A Performance

Teacher participants identified teacher practice as inclusive of classroom instruction and a range of activities that supported teacher and student development. Teacher development activities cited included teacher competitions, seminar attendance, peer observation, and action research projects. Student activities cited included student competitions, in-class and school-wide music and artistic performances, and student presentations of learning. Teachers suggested their participation in student development activities was very important for their teaching performance. A second-grade science teacher attributed school culture and reputation to importance placed on student performance:

> Our reputation really matters so we have to perform high. Some teachers are ok with mediocre performance, but there are not many of these teachers. The culture here is to satisfy students and families and make sure students are happy. Everyone wants to see the students show their learning. The student performance and competitions are very important.

The majority of teachers also explained that their colleagues' and parents' assessment of their teaching was directly shaped by student performance in school activities, competitions, and in assessments. A third-grade physical education teacher explained that because her subject area was "not academic," competitions were given a great deal of emphasis. In contrast, math, Chinese, and English teachers reported student performance on assessments was a core element of teacher performance. One third-grade English teacher explained: "There is a lot of pressure for students to perform well on exams. This is a big part of our teaching. Students need to feel happy, but they still must achieve. There is no choice."

Expectancy, Instrumentality, and Valence

Teachers' perceptions of their practice were directly tied to their views on expectancy, instrumentality, and valence. Overall, perceptions of expectancy were high, and perceptions of

instrumentality and valence were low.

Expectancy. Teachers across all academic domains suggested effort via participation in activities was incentivized. Rewards for activity participation were "obvious," "predictable," and "expected" because bonuses were specifically intended to award increased effort. Physical education and art teachers reported that they were compensated for their effort in district, city, and national student competitions. They suggested these types of activities were particularly important in their disciplines and strong measures of student success. A fifth-grade art teacher explained:

> A lot of effort goes into the competitions. The competitions are very important to our students and for our reputation. It is good we get merit pay for doing well in these competitions because it takes a lot of our time and hard work. We appreciate these bonuses, but the competitions are still very tedious.

Other teachers, particularly the *banzhuren*, all of whom are grade-level heads and teach Chinese, described they received awards for effort in activities related to teacher development. A third-grade Chinese teacher, who was also *banzhuren*, reported that she spent a great deal of time on these professional development activities. She reported that she received about 3000 yuan (430 dollars) for her extra *banzhuren* tasks and her participation in professional development.

Overall, teachers expressed positive expectancy of rewards as associated with activities. When probed for how effort on classroom instruction was rewarded, teachers explained this was simply not the function of incentives. A fourth-grade music teacher explained:

> Merit pay is good for increasing effort in some activities. That is the only thing that merit pay is for. Activities are things like teacher competitions and lesson study and research. A lot of activities are things we do with kids. But merit pay is not for teaching, it is for activities.

This teacher was in the minority in her explicit stance that instruction and bonuses were not tied together. However, most teachers echoed this sentiment. No teacher participant identified an

instance in which their classroom instruction efforts were rewarded.

Instrumentality. While the majority of participants expressed high expectancy (i.e., their effort in teaching and student development activities was rewarded), they were unclear on exactly and how these rewards manifested. Teachers reported that every faculty member received merit pay each month, but the amount varied and was unpredictable. More specifically, they were unclear on how much merit pay they received, how merit pay corresponded with their performance, and why merit pay varied from month to month. Two teachers, who both had several years of experience at M1, were entirely unaware of the merit pay process. When I probed for salary figures, these few teachers replied, "I cannot remember…I'm not sure…my paystub does not tell me."

The other 18 teachers, however, knew their salary included a base salary and merit pay but were unsure of exactly how merit pay worked and what to expect on a monthly basis. A third-grade teacher leader of information technology and art reported his base salary as 515 dollars per month. Providing contradictory information, he reported his merit pay as both a fixed 120 percent of his base salary but also a bonus that changed monthly:

> Oh, I don't know exactly how much. It changes each month. I think my salary is about 3,600 yuan. It's based on your professional title. This is the first part of our salary. The second part varies for different persons. For example, if you are an art teacher, your coefficient is 1, if you are banzhuren, your coefficient is 1.3. When you are vice banzhuren, your coefficient is more than 1, but I can't remember the exact number. I am an art teacher, so my coefficient is 1. But I am also the school art teacher leader, so I get an extra 0.2, so my coefficient is 1.2. 1 multiply by 1.2 is the second part of my salary. But I don't know the merit pay that I got recently.

When probed for examples of how the performance translated into awards, a teacher of first- and second-grade art also narrated a vague, unclear understanding of instrumentality:

> Interviewer: How much merit pay did you receive last year?

Participant: I didn't count. Now our salary increased. In the past, if my coefficient was 1, I only got 1600 (229 dollars), last year I got 1900 (271 dollars), and now I can get 2500 (357 dollars).
Interviewer: Can you explain why this changed?
Participant: I don't know. It just changed.

A second-grade Chinese teacher suggested her salary in September of 2017 was approximately 643 dollars per month, or approximately 7,700 dollars per year. Yet she reported her total salary for the 2015-16 school year was between 8,570 to 10,000 dollars.

Interviewer: Would you mind tell us how much was your total salary last month?
Participant: There was some fluctuation in my salary recently. Some extra money was in it.
Interviewer: what is the extract money?
Participant: There were four holidays. And we got some money because of these holidays. I think my salary is about 4,500 per month (643 dollars).
Interviewer: How much was your total salary last year?
Participant: I think it was about 60,000 or 70,000 (8,570 dollars or 10,000 dollars)

Valence. Despite that most teachers' perceptions of the attainability of performance-pay were vague, and at times inconsistent, the majority of teachers found bonuses slight and insignificant in salary improvement. My co-researcher and I probed several teachers to better understand their low perceptions of instrumentality. We wondered, do teachers, and if not, why don't teachers, try to understand salary breakdowns, the processes through which merit pay is determined, and ways to increase performance pay?

Teachers reported understanding merit pay did not warrant their time because bonuses were nominal. All 20 participants reported their merit pay as between 285 to 715 dollars per month and their base salary as between 430 to 930 dollars per month. The median merit pay reported was 430 dollars, and the median base salary reported was 715 dollars. Twelve of the participants worked as teachers in Shanghai before 2009. All of these teachers reported their salaries increased by no more than three percent since the onset of PBC. Of the eight teachers who had six or less years of experience, six teachers reported their merit pay as approximately

one-third of their total monthly salary. All 20 teachers reported their salaries did not fluctuate by more than 140 dollars per month, or about 10 percent of their total salary.

A fourth-grade Chinese teacher with over 30 years of experience described how both before and after compensation reform in 2009, her salary remained in the 800 dollars per month range. She explained this salary was neither enough for her nor her colleagues:

> I think maybe merit pay should offer more money. These days we all talk about dividing a pie fairly, but the pie is too small. So no matter how you divide it, everyone gets almost the same size of the pie. It doesn't make any difference. Maybe we should enlarge the pie first, and then we can talk about how to divide it.

Another veteran fifth-grade teacher echoed these sentiments. She said, "before merit pay, my pay was weak. After merit pay, my pay is still weak, it just is more complicated and more uncertain."

Fairness. Teachers expressed mixed perceptions on fairness. There was little dispute on the nature and appropriateness of the goals of merit pay. As explained in Chapter 4, teachers expressed merit pay aimed to incentivize teachers to put forth greater effort in teaching and student support activities but had little to do with improved teaching quality. However, participants expressed mixed views on the distribution of merit pay. Half of the 12 veteran teachers expressed disappointment that their salaries and those of new teachers were basically the same. Both the less experienced and veteran teachers took issue with awarding bonuses to "lazy" teachers for two reasons: 1) bonuses did not vary a great deal and thus there was little difference in the rewards for "hardworking" and "lazy" teachers, and 2) bonuses sizes were too insignificant to provoke "lazy" teachers to reform their practices.

M1's PBC policy allowed teachers to appeal to the school administration in the event that a teacher holds concerns with PBC processes and distribution. Several teachers reported they saw little purpose in such appeals because, as mentioned earlier, bonuses were not reflective of teaching quality and were insignificant. In addition, this group of teachers explained that policy

was not being carried out as designed, but the work culture called for discretion and deference to authority. For example, one teacher explained:

> I participated in a research activity. The merit pay policy said I should get 350 (50 dollars) for my participation, but I never received this. I think since my students did not perform well on exams, maybe the school did not want to give me the money…I do not see a purpose in asking why. We don't do that. We don't really talk about merit pay. And also it is very little money.

She and others explained that it was inappropriate to discuss bonuses because the only way to maintain fairness was through unawareness.

Motivation Driven By Values Of Self-Efficacy and Peer Respect

As teachers discussed their perspective son pay, they explained their motivation to improve their teaching performance was driven by appreciation for public recognition and a sense of individual responsibility for the collective mission of student success. Even though teachers were not clear on whether and how they were awarded merit pay, they participated in lesson study, action research, and teaching competitions because they knew this would improve their teaching and because, as a fifth-grade art teacher explained, it was "part of the culture" and "just what we do here." For the teacher leader participants, engagement in professional development activities was worthy job requirement because it made them "proud" of their work. A physical education teacher elaborated:

> At first, I thought a bonus meant that my work is recognized, and my ability is recognized. The bonus enhanced my enthusiasm about my job. When we didn't have merit pay, everyone received the same salary whether he did his work or not. Then schools were in a mess, and the quality of the school was not improved. The school needs this policy to develop. I think bonuses means something to me. It makes me devote full attention to my work. The bonus is not much, but it means that I am being recognized.

This teacher's sentiments mirrored those of her peers. When we discussed the benefits and positive influences of merit pay, the majority of participants perceived merit pay as a vehicle for recognition for individual work and recognition for team success.

In addition, teachers reported they enjoyed opportunities to receive certificates and be recognized in public by their peers, and especially their principals and researchers from universities. The majority of teachers explained they were moved by opportunities to earn certificates, model lessons to large groups of teachers, positive parent feedback, and positive principal feedback. The "extra" subject teachers of physical education, art, and music explained they felt a great deal of pressure to help students win competitions but "really enjoyed" the respect from fellow teachers and parents from success in competitions. A first- and second-grade music teacher elaborated: "It is a lot of work for us to take students to these competitions. There is a lot of pressure to win. But the children feel so happy. Their parents are so happy. When they are happy, they value us. We want their respect."

Discussion

The M1 teacher participants reported they were unsure of the exact aims of merit pay, yet they perceived bonuses intended to strengthen teacher enthusiasm, improve job satisfaction, and incentivize teacher participation activities that supported teacher and student development. However, about half of teachers suggested there was no relationship between merit pay and teacher quality. These teachers perceived merit pay had either no effect or a positive effect on their own and their colleagues' enthusiasm and professional satisfaction, was likely to provoke poorly performing or "lazy" teachers to engage in more activities, and offered a deserved salary boost to "good" and highly engaged teachers. Yet this group was quite clear that teachers' increased school engagement did not necessarily lead to improved quality.

In the view of policymakers, student activities, research, and professional development served as proxies and stimuli for teacher quality. In other words, rewards lead to greater extrinsic effort, which generates improved teaching quality. The teachers in this study, however, did not

draw such a connection. Nearly all teacher participants reported they did not see a connection between participation in student and teaching activities and teacher quality, and approximately half of the participants reported that there was no relationship between merit pay and teaching quality. Bonuses thus functioned as salary supplements rather than incentives to (re)contour one's quality of teaching. Relatively harmonious teacher relationships withstood and axed the potential competitive outcomes of incentivism. In teachers' words, at best, bonuses incentivized increased effort and work time, and this increased effort differentiated high from low performing teachers but did not dramatically shape the improvement trajectory of low performing teachers.

These patterns speak to two of Goldhaber and colleagues' (2008) propositions. In School M1, teachers received more pay than they did prior to 2009, when merit pay was not institutionalized. Also, M1 constructed a PBC policy in a way that afforded the school greater knowledge about teaching in order to award bonuses. However, when teachers spoke of rewards for effort, they referred to participation in activities related to teaching and student development. When they spoke of performance, they referred to participation in activities *and* instruction. Teachers were clear that their *effort* was rewarded, but they alluded, and in some cases were quite certain, their performance was not rewarded. That is, teachers drew a clear distinction between effort and performance.

The cognitive frame that has propelled incentivism and performance-based teaching into the global education policy sphere presumes a neat, direct link between financial incentives and teacher performance. This idea undergirds Shanghai educational leaders' approach to performance-based compensation for teachers. Shanghai city and district policymakers, as well as the M1 administration, adopted PBC under the premise that financial bonuses would improve teacher quality. The city and district mandated schools to develop performance assessment and

pay standards and procedures that included teacher self-, peer-, and committee-based performance assessment of teacher performance. Merit pay was to be allocated selectively and awarded to the top tier of teachers and determined based upon teacher title; ability; achievements; and progress in moral education, research, and professional development. School M1's vision for merit pay and procedures were developed under these aspirations.

The "front-line workers" (OECD, 2014) in this study uncovered a key breakdown in incentivism: the behaviors that link rewards to quality are ill defined and may have little to do with a teacher's extra or more of the same work. The question of 'what is quality' remains highly contested in global education policy, let alone in scholarship on teacher preparation, development, certification, and pedagogy (Darling-Hammond, 2013). In GEP, notions of quality are generic and responsive to the needs of globalization and innovation. In *Surpassing Shanghai*, Tucker (2011) asserted high quality teachers hold "a high level of general intelligence, a solid mastery of the subjects to be taught, and a demonstrated aptitude for engaging students and helping them understand what is being taught" (p. 178). From a relational perspective on ideas and institutional change, the nebulous, innovation and intelligence-centered character of 'quality' translates well into the normative ideology of globalization and innovation. Because quality needs to remain nimble to the unforeseen and ever-changing demands of globalization and innovation, it retains a vague yet convincing character.

These findings directly interrogate the economic persuasion of incentivism from two angles. First, advocates have posited that extrinsic rewards, such as financial bonuses, unequivocally provoke teachers to exert their maximum effort. These data dispute this claim. Most participants suggested they were clear bonuses were *intended* to incentive greater effort, but because the valence of rewards was unattractive, bonuses bore little impact on their effort.

Relatedly, some teachers explained they exerted their maximum effort in the name of collective responsibility and in pursuit of positive public recognition. Secondly, advocates have posited that teachers' motivation to exert maximum effort leads to better individual performance and an improved teacher workforce, particularly because low performing teachers exit. Yet these participants suggested the effort, whether maximum or not, that was rewarded (by policy design) did not necessarily bear upon instruction, a core function of teacher quality. In addition, participants were quite clear that "poor" or "lazy" teachers were not incentivized to change their work because, again, valence was unattractive.

To better understand what might be necessary to garner the improved individual teachers and an improved workforce outcomes espoused by incentives advocates, these participants' perspectives on expectancy, instrumentality, valence, and fairness are instructive. Teachers held low perceptions of valence and instrumentality. They were unclear as to whether and how their performance led to bonuses. Most teachers found these bonuses insignificant and insufficient and thus did not take action to better understand bonus distribution. Yet teachers also held high perceptions of expectancy. They expressed confidently that bonuses were firmly linked to teacher effort, but mostly the kind that took place outside of the classroom.

In their examination of the potential of a voluntary incentives program to motivate changes in teacher practice, Rice and colleagues (2015) concluded the valence of reward was critical:

> The vast majority of our survey respondents indicated that they increased their effort, at least in the short term, as a result of the program—perhaps because the budget constraints of the district made the rewards particularly salient to educators as a rare opportunity to augment their salaries, and perhaps because the size of the maximum potential awards available through the FIRST program met Lawler's (1990) recommendation that the incentives should constitute at least 10% to 15% of one's salary….teacher reactions are, at least in part, a function of the size of the rewards…success might depend on how payout amounts and responses to payouts evolve over time. (p. 45)

This investigation found that a ten to 15 percent bonus did not elicit high perspectives on valence. Potential rewards at School M1 were often well above the recommended ten to 15 percent. However, participants perceived their monthly merit pay augmented their salary in more of a *supplement* fashion than as a bonus, and these supplements were insufficient to meet their increasing monthly expenses in a city that was rapidly becoming costlier.

Rice and colleagues' (2015) research further found empirical evidence that teachers held more positive perspectives on bonuses that were seamlessly tied with effort and performance:

> Expectancy theory, along with the concept of attainability from goal setting theory, predicts that educators are likely to be more supportive of incentives that have dependable connections between effort, performance, and reward. Empirical evidence from our study supports this idea. Teachers in our study were most positive about the payouts they viewed as attainable and dependable; they were most negative about the payouts that they perceived to be beyond their control regardless of the amount. (p. 45)

In line with this work, the findings in this study suggest teachers were keenly aware of the connections between merit pay policy and effort. Teachers perceived the policy divorced bonuses from performance. From a theory of agency perspective, it is possible that policymakers, at all levels, were stumped and/or found it too tedious to gather information about teacher performance. Instead, attaching bonuses to specific activities proved a standardized, efficient method of gaining information about, exacting teacher effort, and holding mass amounts of teachers accountable to the same standards for bonuses.

CHAPTER 6

CONCLUSIONS AND IMPLICATIONS

Recently, OECD Director Andreas Schleicher extolled both the importance of teaching quality, the quality of Shanghai's teachers, and the value of teaching for innovation:

> …And nowhere does the quality of a school system exceed the quality of its teachers. The East Asian school systems all pay great attention to how they select and train their staff. And when deciding where to invest, they prioritize the quality of teachers over the size of classes. They provide intelligent pathways for teachers to grow in their careers. High-performing countries have also moved on from bureaucratic control and accountability to professional forms of work organization. They encourage their teachers to make innovations in pedagogy, to improve their own performance and that of their colleagues, and to pursue professional development that leads to stronger education practice. The goal of the past was standardization and compliance; but today's top-performing countries value inventiveness. In the past, policy focused on providing education; today's top school systems focus on outcomes, shifting from looking upwards in the bureaucracy to looking outwards to the next teacher, the next school, to create networks of innovation. You can see that nowhere better than in Finland or Shanghai.

His sentiments underscore Shanghai's increasing base of admirers and the potential for this high performing system to inform policy adoption in the global education policy sphere. While this study did not squarely study innovation, incentivism is one particular policy device that Shanghai policymakers levy in the name of innovation. The claim with which incentivism has penetrated high-stakes teacher accountability in the global sphere is fairly straightforward and compelling. Inherent in the theory of action behind PBC is the idea that bonuses motivate educators to behave in desirable ways, incentivize strong teachers to join and stay in the workforce, and encourage weak teachers to exit (Hanushek & Lindseth, 2009; Weiner, 1980). The idea that an incentive renovates teaching, learning, schools, and markets is a policy proposal that is well received by policy actors, such as those in Shanghai, who are persuaded by the current global education reform framework of innovation, competition, and globalization.

Shanghai found great promise in the power of incentives to improve teaching quality. They pushed school districts and leaders to create comprehensive performance appraisal and pay plans that would uplift teacher quality and transform the composition of the teacher workforce in a positive direction. For this to happen, the incentives structure would need to directly impact classroom instruction and identify and persuade poor instructors to improve at an appropriate pace or exit the school. Instead, the incentive structure awarded teachers largely for more of the same work. In a culture of stability, equality, and team-based professional norms, performance incentives became bland. In the eyes of the M1 teachers, bonuses were salary add-ons, and at their best, these supplements were insufficient to meet their growing daily economic needs.

Johnson (2004) suggested incentives advocates misunderstand the bedrocks of teachers' motivation: "In itself, higher pay is unlikely to retain teachers – particularly the most able among them – if they cannot attain the intrinsic rewards for which they initially entered teaching" (p. 46). While others seek to optimize performance and recompose the teacher workforce, teachers seek intrinsic rewards such as trust, team membership, community membership, professional recognition, job security, and a viable professional future that continues to build on these rewards. In addition to the high political and policy costs and undesirable effects of incentives, this study suggests incentives have the potential to further direct teacher attention away from instruction and toward ad hoc activities. Indeed, professional development and student development activities are a core part of schooling and speak to teacher performance. Yet for incentives to achieve the desired end result of improved student outcomes, instruction must be a core focus. These findings build on previous research that suggests teachers are inspired to look inward and analyze their practice in the presence of opportunities for recognition and encouragement from their superiors, peers, parents, and students.

Implications

Although this particular study took place in Shanghai, it is also a story about incentivist reforms more broadly. My findings are particular to School M1, but they have important implications for policy, for the leaders of other education systems that are in the process of adopting or considering the adoption of performance-based compensation, as well as implications for research.

Policymakers often draw quite loosely from economic concepts to inform their theories about issues such as accountability. As equity concerns rise, policymakers default to models of standardization and efficiency to affect change in timely, measurable manners. But policymakers must also clarify how incentivism is expected to operate in public schools, how its operation differs in private versus public markets, and its potential effects. Policymakers must also consider the potential outcomes of financial incentives, both the positive and the negative, by paying particular attention to what drives instructional and school improvement. For example, incentivizing the achievement of sales targets in private firms is quite normal, but do policymakers expect to devote public dollars to incentive teaching quality in the same fashion? If so, these findings provoke policymakers to consider the ways in which School M1 teacher participants' perspectives interrogated the underpinnings of incentivism. When we zoom in on incentivism in the classroom and consider how teachers take up policies targeted at their improvement, these questions push theory and policy on incentivism to revise the currently thin and seemingly erroneous ideology of the linear link between money and motivation in the context of schools.

My findings also implore research, leadership, and policy to place a clear emphasis on building positive school culture that is laser focused on teamwork. The culture of School M1 is

one that nurtures Confucian values of individual responsibility for collective success and equality. In doing so, the school unites rather than pits teachers against one another and capitalizes on rather than manipulating harmony. Norms of deference to authority also serve to silence teacher concerns about the fairness and efficacy of incentives. School leaders and policymakers should consider how positive school culture may be a more efficacious route than incentivism to teaching and school improvement. Yet leaders and policymakers must also consider how to uplift rather than silence teacher voice. Since perspectives of pay mediate the enactment of compensation reform, nurturing and honoring teacher voice, particularly in an additive and assets-based team atmosphere, may begin to thrust teacher accountability policy in a positive direction.

The M1 teachers implore policymakers and researchers alike to consider: What is the efficacy of bonuses when policymakers conflate effort with quality? What exact behaviors should policy incentivize to boost quality? If bonuses neither incentivize low performing teachers to perform well nor force low performing teachers to exit the workforce, what becomes of the incentivist policy paradigm? The M1 policy took incentives out of the classroom, and thus teachers equated bonuses more closely with effort in student development activities rather than with the core activity of teachers – teaching. When bonuses direct attention away from instruction, interrupt improvement processes are interrupted. Moreover, leaders must anticipate the consequences of a bland policy that incentivizes effort that has already been normalized in an equality-centered school culture and awards bonuses that are perceived as low and insufficient. The participant reactions in M1 to bland incentivism suggest one particularly grave consequence is apathy and ambivalence toward incentives.

Perhaps most importantly, policymakers and leaders have an opportunity to consider how to build an incentive structure that moves away from the high policy and political costs of VAMs and of public rankings of teacher performance. Student performance is just one of several performance metrics in School M1, and it is one that teachers say holds little weight in performance appraisal and pay. As the U.S. and other Western countries ramp up test-based accountability and tie student performance metrics to evidence-based practice and policy, teacher evaluation, and performance-based compensation, Shanghai and School M1 provide a blueprint to reenvision how incentives provoke a renewed focus on students and teaching rather than on testing.

Future Research

The findings from this study suggest at least three directions for future research. First, my research indicates that we need better proxies for incentives to study its effects. My work begins to extend and elaborate theory to illuminate the black box of incentivism in schools, and there is a need to explore this process in other contexts. The context of Shanghai paves a new road for examining incentivism in high performing, influential systems with Confucian cultural norms. By looking at incentivism in systems that fit this road, such as Singapore, South Korea, and Japan, we can begin to illuminate additional perspectives of teachers and microinstitutional variables that mediate policy enactment. Second, my methodological approach privileges the perspectives of teachers, which are sorely needed in research on incentivism. Additional mixed methods research that builds on interview data to develop survey instruments will afford an opportunity to isolate the specific behaviors of teachers that are shaped by incentives. Such rich, nuanced analyses that focus on and triangulate the perspective of teachers can lead us to better policies for teacher improvement. Mixed methods investigations will help zero in on the

variables of teaching quality at which to target incentives. Contextual, theory-laden investigations of incentives in the classroom can continue to zero in closely on the behaviors that ensue when teachers are faced with incentives. Third, future research will benefit from a closer look at the role of teacher leaders in the enactment of incentivism and their understandings of instructional improvement processes. As teacher leadership gains potency in school improvement planning, understanding the role of these individuals in brokering teacher interpretation of policy is of great importance.

As leaders take up and enact incentivism upon their teachers and schools, they take a potentially harmful approach of linking improvement with teachers' sense of confidence and fear. Research would benefit from multi method and comparative studies that focus on the cultural foundations of schools and school improvement. By seeing incentivist policy as situated in the sociocultural and political environments of schools in diverse contexts, we can continue to better assess the prevailing hypothesis that there is a neat relationship between an incentive, teacher practice, and student learning.

REFERENCES

Al-Amer, R., Ramjan, L., Glew, P., Darwish, M., & Salamonson, Y. (2014). Translation of interviews from a source language to a target language: examining issues in cross-cultural health care research. *Journal of Clinical Nursing, 24,* 1151–1162, doi: 10.1111/jocn.12681

Amrein-Beardsley, A., Polasky, S., & Holloway-Libell, J. (2016). Validating "value added" in the primary grades: One district's attempts to increase fairness and inclusivity in its teacher evaluation system. *Education Assessment Evaluation Association, 28,* 139–159. doi: 10.1007/s11092-015-9234-5

Au, W. & Lubienski, C. (2016). The role of the Gates Foundation and the philanthropic sector in shaping the emerging education market: Lessons from the US on privatization of schools and education governance. In A. Verger, C. Lubienski, & G. Steiner-Khamsi (Eds.), *World Yearbook of Education 2016: The Global Education Industry* (27-43). New York, NY: Routledge.

Balch, R., & Springer, M. G. (2015). Performance pay, test scores, and student learning objectives. *Economics of Education Review, 44,* 114-125. doi: 10.1016/j.econedurev.2014.11.002

Ball, S. J., Maguire, M., & Braun A. (2012). How schools do policy: Policy enactments in secondary schools. New York, NY: Routledge.

Ballou, D., & Podgursky, M. (2003). Teacher attitudes toward merit pay: Examining conventional wisdom. *Industrial and Labor Relations Review, 47*(1), 50-61.

Baker, G. P., Jensen, M. C., & Murphy, K. J. (1988). Compensation and incentives: Practice vs. theory. *Journal of Finance, 43*(3), 593-616.

Berliner, D. (2015). The Many Facets of PISA. *Teachers College Record, 117*(1), 1-20. Retrieved from http://www.tcrecord.org ID Number: 17726

Berliner, D. C., & Glass, G. V. (Eds.). (2014). *50 Myths and lies that threaten America's public schools: The real crisis in education*. New York, NY: Teachers College Press.

Berman, R. C. & Tyyskä, V. (2010). A Critical Reflection on the Use of Translators/Interpreters in a Qualitative Cross-Language Research Project. *International Journal of Qualitative Methods, 10*(1).

Bozionelos, N., & Wang, L. (2007). An investigation on the attitudes of Chinese workers towards individually based performance-related reward systems. International Journal of Human Resource Management, 18(2), 284-302. doi:10.1080/09585190601102497

Braun, A., Ball, S., Maguire, M., & Hoskins, K. (2011). Taking context seriously: Towards explaining policy enactments in the secondary school. Discourse: Studies in the Cultural Politics of Education, 32(4), 585-596. doi: 10.1080/01596306.2011.601555

Brodsky, A., DeCesare, D., & Kramer-Wine, J. (2010). Design and implementation considerations for alternative teacher compensation programs. *Theory Into Practice, 49*, 213–222. doi: 10.1080/00405841.2010.487757

Bryan, J. F., & Locke, E. A. (1967). Goal-setting as a means of increasing motivation. *Journal of Applied Psychology, 51,* 274–277.

Bryman, A. (2004). *Social research methods.* Oxford, England: Oxford University Press.

Burningham, G. (2014). Lessons from the world's best public school. *Newsweek*. Retrieved from file:///Users/priyagoel/Box%20Sync/diss/data/documents/Burningham_2014_Lessons%20From%20the%20World's%20Best%20Public%20School.htm

Chiang, H., Wellington, A., Hallgren, K., Speroni, C., Herrmann, M., Glazerman, S., & Constantine, J. (2015). Evaluation of the Teacher Incentive Fund: Implementation and impacts of pay-for-performance after two years. *National Center for Education Evaluation and Regional Assistance*. Retrieved from http://ies.ed.gov/ncee/pubs/20154020/pdf/20154020.pdf

Chubb, J. E., & Moe, T. M. (1990). *Politics, Markets, and America's Schools*. Washington, DC: Brookings Institution.

Cissell, G. (2010). Kentucky and education reform: The issue of pay-for-performance. *Journal Of Law & Education*, *39*(1), 119-127.

Clees, W. J., & Nabors, M. L. (1992). Teacher incentive programs: Do they make better teachers?. (1992). *Education*, *113*(1), 145-148.

Coughlan, S. (2014, April 10). Highest ever number of school staff. *BBC News.* Retrieved from http://www.bbc.com/news/education-26973916

Creswell, J. (2003). Research design: Qualitative, quantitative, and mixed methods approaches (2nd Ed.). London, New Delhi: Sage Publications.

Darling-Hammond, L. (2013). *Getting teacher evaluation right: What really matters for effectiveness and improvement*. New York, NY: Teachers College Press.

Deal, T. E., & Peterson, K. D. (2016). Shaping school culture. San Francisco, CA: John Wiley & Sons.

Earn, B. M. (1982). Intrinsic motivation as a function of extrinsic financial rewards and subjects' locus of control. Journal of Personality, 50(3), 360. doi:10.1111/1467-6494.ep7379391

Eberts, R., Hollenbeck, K., & Stone, J. (2002). Teacher performance incentives and student outcomes. *Journal of Human Resources, 37*(4), 913-927.

Edwards, Rosalind (1998). A critical examination of the use of interpreters in the qualitative research process. *Journal of Ethnic and Migration Studies, 24*(2), 197-208.

Elliott, S. W., & Hout, M. (Eds.). (2011). Incentives and test-based accountability in education. National Academies Press. Washington, D.C.

Esposito, N. (2001). From meaning to meaning. Qualitative Health Research, 11, 568-579.

Evans, L. (2011). The 'shape' of teacher professionalism in England: Professional standards, performance management, professional development and the changes proposed in the 2010 White Paper. *British Educational Research Journal, 37*(5), 851-870. doi: 10.1080/01411926.2011.607231

Elpus, K. 2011. Merit pay and the music teacher. Arts Education Policy Review 112 (4): 180–90.

EWC Annual Statistics Digest (2014). Retrieved from http://www.ewc.wales/site/images/documents/statistics/GTCW/An_Analysis_of_registered_Supply_Teachers_2014.pdf

Fersch, B. (2013). Meaning: lost, found or 'made' in translation? A hermeneutical approach to cross-language interview research. *Qualitative Studies, 4*(2), 86-99.

Flick, U. (1998). *An introduction to qualitative research.* London, England: Sage.

Fong, S. C. L., & Shaffer, M. A. (2003). The dimensionality and determinants of pay satisfaction: A cross-cultural investigation of a group incentive plan. *International Journal of Human Resource Management, 14*(4), 559-580.

Fontana, A., & Frey, J. (2000). The interview: From structured questions to negotiated text. In N. K. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed.), 61-106. Thousand Oaks, CA: Sage.

Garton, S. & Copland, F. (2010). "I like this interview; I get cakes and cats": The effect of prior

relationships on interview talk. *Qualitative Research, 10*(5), 533-551.
Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory.* Chicago, IL: Aldine Publishing Company.

Glewwe, P. N., & Kremer, M. (2010). Teacher incentives. *American Economic Journal: Applied Economics, 2*, 205–227.
Goldhaber, D. (2002). Teacher quality and teacher pay structure: What do we know, and what are the options?. *The Georgetown Public Policy Review, 7,* 81-93.
Goldhaber, D., DeArmond, M., Player, D., & Choi, H. (2008). Why do so few public school districts use merit pay? *Journal of Education Finance, 33*(3), 262-289.
Goldhaber, D., DeArmond, M., & Deburgomaster, S. (2011). Teacher attitudes about compensation reform: Implications for reform implementation. *Industrial and Labor Relations Review, 64*(3), 441-463.
Goodman, S. F., & Turner, L. J. (2013). The design of teacher incentive pay and educational outcomes: Evidence from the New York City bonus program. *Journal of Labor Economics, 31*(2), 409-420. Retrieved from http://www.jstor.org/stable/10.1086/668676
Gruenert, S., & Whitaker, T. (2015) School culture rewired: How to define, assess, and transform it. Association for Supervision and Curriculum Development.
Hallinger, P., Heck, R. H., & Murphy, J. (2014). Teacher evaluation and school improvement: An analysis of the evidence. *Educational Assessment, Evaluation and Accountability*, *26*(1), 5-28. doi: 10.1007/s11092-013-9179-5
Haynes, G., Wragg, C., Wragg, T., & Chamberlin, R. (2003a). Threshold assessment: The experiences of teachers who were unsuccessful in crossing the threshold. *Research Papers in Education, 18*(1), 25-44. doi: 10.1080/0267152032000048569
Haynes G., Wragg, C., Wragg, T., & Chamberlin, R. (2003b). Performance management for teachers: Headteachers' perspectives. *School Leadership & Management: Formerly School Organisation, 23*(1), 75-89. doi: 10.1080/1363243032000080041
Hanushek, E. A., & Lindseth, A. A. (2009). *Schoolhouses, courthouses, and statehouses: Solving the funding-achievement puzzle in America's public schools.* Princeton, NJ: Princeton University Press.
Heneman, H. G. III, Milanowski, A., & Kimball, S. (2007, February). Teacher performance pay: Synthesis of plans, research, and guidelines for practice (Policy Brief No. 46). Philadelphia, PA: Consortium for Policy Research in Education. Retrieved from www.cpre.org/images/stories/cpre_pdfs/RB46.pdf
Heneman, H. G., III. (1998). Assessment of the motivational reactions of teachers to a school-based performance award program. *Journal of Personnel Evaluation in Education, 12,* 43–59.
Henig, J. (2013). *The end of exceptionalism in American education: The changing politics of school reform.* Cambridge, MA: Harvard Education Press.
Higgins-D'Alessandro, A., & Sadh, D. (1997). The dimensions and measurement of school culture: Understanding school culture as the basis for school reform. *International Journal of Educational Research, 27,* 553–69.
Hofstede, G. & Minkov, M. (2010). Long- versus short-term orientation: new perspectives. Asia Pacific Business Review, 16(4), 493-504. doi: 10.1080/13602381003637609
Inhetveen, K. (2012). Translation Challenges: Qualitative interviewing in a multi-lingual field. *Qualitative Sociology Review, 8*(2), 28-45. Retrieved from

http://www.qualitativesociologyreview.org/ENG/archive_eng.php
Jagosh, J. & Boudreau, D. (2009). Lost and found in translation: An ecological approach to bilingual research methodology. *International Journal of Methods 8,* 102–114.
Jennings, J. L. (2012). The effects of accountability system design on teachers' use of test score data. *Teachers College Record*, *114*(11), 1-23.
Johnson, S. M. (2004). Paralysis or possibility: What do teachers unions and collective bargaining bring? In R. Henderson, W. Urban, & P. Wolman (Eds.), *Teachers unions and education policy: Retrenchment or reform?* (1-30). Oxford, United Kingdom: Elsevier.
Kelley, C., Conley, S., & Kimball, S. (2000). Payment for results: Effects of the Kentucky and Maryland group-based performance award programs. *Peabody Journal of Education*, *75*(4), 159-199. doi: 10.1207/ S15327930PJE7504_8
Kobakhidze, M. N. (2010). Teacher incentives and the future of merit-based pay in Georgia. *European Education, 42*(3), 68–89. doi: 10.2753/EUE1056-4934420304
Kreps, D. M. (1997). Intrinsic motivation and extrinsic incentives. *American Economic Review, 87*(2), 359-364. Retrieved from http://www.jstor.org/stable/2950946
Labaree, D. F. (2014). Let's measure what no one teaches: PISA, NCLB, and the shrinking aims of education. *Teachers College Record, 116*(9), 1-14. Retrieved from http://www.tcrecord.org.proxy2.library.illinois.edu/library/Issue.asp?volyear=2014&number=9&volume=116
Lai, M., & Lo, L. N. K. (2007). Teacher professionalism in educational reform: The experiences of Hong Kong and shanghai. *Compare: A Journal of Comparative Education, 37*(1), 53-68. doi:10.1080/03057920601061786
Larkin, P. J., Dierckx de Casterlé, B., & Schotsmans, P. (2007). Multilingual translation issues in qualitative research: Reflections on a metaphorical process. *Qualitative Health Research, 17*(4), 468-476. doi: 10.1177/1049732307299258
Lavy, V. (2007). Using performance-based pay to improve the quality of teachers. *Future Of Children*, *17*(1), 87-109.
Lawler, E. E. (1976). Participation and pay. *International Journal of Production Research, 14*(3), 367.
Lawler, E. E., III. (1983). *Pay and organization development.* Reading, Mass: Addison-Wesley.
Lazear, E. (2003). Teacher Incentives. *Swedish Economic Policy Review, 10,* 197–213.
Leech, N. L. & Onwuegbuzie, A. J. (2007). An array of qualitative data analysis tools: A call for data analysis triangulation. *School Psychology Quarterly, 22*(4), 557–584. doi: 10.1037/1045-3830.22.4.557
Leigh, A. (2012). The economics and politics of teacher merit pay. *CESIfo Economic Studies, 59*, 1-33. doi: 10.1093/cesifo/ifs007
Locke, E. A., & Latham, G. P. (2006). New directions in goal setting theory. *Current Directions in Psychological Science, 15,* 265–268.
Lubienski, C., Scott, J., & DeBray, E. (2011). The rise of intermediary organizations in knowledge production, advocacy, and educational policy. *Teachers College Record*. Retrieved from http://www.tcrecord.org ID Number: 16487
Lubienski, C. A., & Lubienski, S. T. (2014). *The public school advantage: Why public schools outperform private schools.* Chicago, IL: University of Chicago Press.
Lubienski, C., Brewer, J., & La Londe, P. G. (2015). Orchestrating policy ideas: Philanthropies and think tanks in US education policy advocacy networks. *Australian Education Researcher, 43*(1), 55-73. doi: 10.1007/s13384-015-0187-y

Lussier, D. F., & Forgione, P. D. (2010). Supporting and rewarding accomplished teaching: Insights from Austin, Texas. *Theory Into Practice, 49*, 233–242. doi: 10.1080/00405841.2010.487771

Luthans, F., & Stajkovic, A. D. (1999). Reinforce for performance: The need to go beyond pay and even rewards. *Academy of Management Executive, 13*(2), 49-57. doi:10.5465/AME.1999.1899548

Mahony, P., Hextall, I., & Menter, I. (2002). Threshold assessment: Another peculiarity of the English or more McDonaldisation?, *International Studies in Sociology of Education, 12*(2), 145-168. doi: 10.1080/09620210200200088

Malen, B., Rice, J. K., Matlach, K. B., Bowsher, A., Hoyer, K. M., & Hyde, L. H. (2015). Developing organizational capacity for implementing complex education reform initiatives: Insights from a multiyear study of a Teacher Incentive Fund program. *Educational Administration Quarterly, 5*(1), 133-176. doi: 10.1177/0013161X14522482

Marsden, D. (2014). Teachers and performance pay in 2014: First results of a survey. Centre for Economic Performance Working Paper # 2140.

Marsh, J. A., Springer, M. G., McCaffrey, D. F., Yuan, K., & Epstein, S. (2011). *A big apple for educators: New York City's experiment with schoolwide performance bonuses: Final evaluation report*. Rand Corporation. Retrieved from http://www.rand.org/content/dam/rand/pubs/monographs/2011/RAND_MG1114.pdf

Matteson, S. M. & Lincoln, Y. S. (2009). Using multiple interviewers in qualitative research studies: The influence of Ethic of Care behaviors in research interview settings. Qualitative Inquiry, (15) 4, 659-674. doi: 10.1177/1077800408330233

Matthes, W. A., & Tollerud, T. R. (1990). Teacher beliefs regarding career ladders and merit pay. *Education*, *110*(4), 535-540.

Mercer, J. (2006). The challenges of insider research in educational institutions: Wielding a double-edged sword and resolving delicate dilemmas. *Oxford Review of Education, 33*(1), 1-17.

Meyer, H. D., Tröhler, D., Labaree, D. F., & Hutt, E. L. (2014). Accountability: Antecedents, power, and processes. *Teachers College Record, 116*(9), 1-12. Retrieved from http://www.tcrecord.org.proxy2.library.illinois.edu/library/Issue.asp?volyear=2014&number=9&volume=116

Meyer, H. D. (2013). OECD's PISA: A tale of flaws and hubris. *Teachers College Record*, 1-6. Retrieved from http://www.tcrecord.org ID Number: 17371

Miles, M. B., & Huberman, M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Morrison, N. (2014). What we can learn from the success of Shanghai's schools. *Forbes*. Retrieved from https://www.forbes.com/sites/nickmorrison/2014/02/05/what-we-can-learn-from-the-success-of-shanghais-schools/#3d6eb27c103e

Muralidharan, K., & Sundararaman, V. (2011). Teacher performance pay: Experimental evidence from India. *Journal of Political Economy*, *119*(1), 39-77.

Murnane, R. J., & Cohen, D. K. (1986) Merit pay and the evaluation problem: Why most merit pay plans fail and a few survive. *Harvard Educational Review, 56*(1), 1-17.

Neal, D. (2009). Designing incentive systems for schools. In M. D. Springer (Ed.), *Performance incentives: Their growing impact on American K-12 education* (149-170). Washington D.C.: The Brookings Institution.

OECD. (2009). Design components of incentive pay programmes in the education sector. In

Evaluating and rewarding the quality of teachers: International practices. OECD Publishing. doi:10.1787/9789264034358-5-en
OECD. (2009). Whose incentives? Whose stimuli? Performance management in England. In Evaluating and rewarding the quality of teachers: International practices (pp. 103–119). Paris, France: OECD Publishing. doi: http://dx.doi.org/10.1787/9789264034358-6-en
OECD. (2012a). Shanghai and Hong Kong: Learning to learn. In Lessons from PISA for Japan (pp. 149–177). Paris, France: OECD Publishing. doi: http://dx.doi.org/10.1787/9789264118539-8-en
OECD. (2012b). Policies and practices governing public sector compensation planning. In Public sector compensation in times of austerity (pp. 109–128). Paris, France: OECD Publishing. doi: http://dx.doi.org/10.1787/9789264177758-7-en
OECD. (2014). *TALIS 2013 Results: An international perspective on teaching and learning.* Retrieved from http://www.oecd-ilibrary.org/education/talis-2013-results_9789264196261-en
Odden, A., & Kelley, C. (1997). *Paying teachers for what they know and do.* Thousand Oaks, CA: Corwin.
Odden, A., & Wallace, M. J. (2007). Rewarding Teacher Excellence: A Teacher Compensation Handbook for State and Legal Policy Makers. Consortium for Policy Research in Education, Wisconsin Center for Education Research, University of Wisconsin.
Orland, M. (2015). Research and policy perspectives on data-based decision making in education. *Teachers College Record, 117*, 1-10.
Park, S. & Lunt, N. (2015). Confucianism and qualitative interviewing: Working Seoul to Soul. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 16(2), Art. 7, Retrieved from http://nbn-resolving.de/urn:nbn:de:0114-fqs150274
Patton, M. (1990). *Qualitative evaluation and research methods.* Newbury Park, CA, and London, UK: Sage Publications.
Podgursky, M., & Springer, M. G. (2007a). Credentials Versus Performance: Review of the Teacher Merit pay Research. *Peabody Journal Of Education, 82*(4), 551-573. doi:10.1080/01619560701602934
Podgursky, M. J., & Springer, M. G. (2007b). Teacher Performance Pay: A Review. *Journal Of Policy Analysis & Management*, *26*(4), 909-949.
Programme for International Student Assessment. *PISA 2012 Results.* Retrieved from http://www.oecd.org/pisa/keyfindings/pisa-2012-results.htm

Programme for International Student Assessment. *PISA 2012 Results.* Retrieved from http://www.oecd.org/pisa/keyfindings/pisa-2012-results.htm
Race to the Top Act of 2011, S. 844–112th Congress. (2011). Retrieved from http://www.govtrack.us/congress/bills/112/s844
Raham, H. (2000) Cooperative Performance Incentive Plans. *Peabody Journal of Education*, *75*(4), 142-158. doi: 10.1207/S15327930PJE7504_7
Ratcliffe, R. (2013, November 5). Inside performance related pay: How will the new system work? *The Guardian.* Retrieved from http://www.theguardian.com/teacher-network/teacher-blog/2013/nov/05/performance-related-pay-new-system-work
Rice, J. K., Malen, B., Baumann, P., Chen, E., Dougherty, A., Hyde, L., Jackson, C., Jacobsen, R., & McKithen, C. (2012). The persistent problems and confounding challenges of

educator incentives: The case of TIF in Prince George's County, Maryland. *Education Policy, 26*(6), 892-933. doi: 10.1177/0895904812465708
Rice, J. K., Malen, B., Jackson, C., & Hoyer, K. M. (2015). Time to pay up: Analyzing the motivational potential of financial awards in a TIF program. *Educational Evaluation and Policy Analysis, 37*(1), 29-49. doi: 10.3102/0162373714524622
Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73,* 417–458.
Robbins, P. & Alvy, H. B. (2004). *The new principal's fieldbook: Strategies for success*. Association for Supervision and Curriculum Development.
Rosenblatt, P. C. (2012). One interviewer versus several: Modernist and postmodernist perspectives in qualitative family interviewing. *Journal of Family Theory & Review, 4,* 96–104. doi: 10.1111/j.1756-2589.2012.00120.x
Rothstein, R. (2004). *Class and schools: Using social, economic, and educational reform to close the achievement gap*. Washington, DC: Economic Policy Institute.
Sahlberg, P. (2011). Finnish lessons. New York: Teachers College Press.
Sahlberg, P. (2015). *Finnish Lessons 2.0*. New York, NY: Teachers College Press.
Schleicher, A. (2016). Opinion: What Asian schools can teach the rest of us. *CNN*. Retrieved from http://edition.cnn.com/2016/12/06/opinions/education-pisa-rankings-china/index.html
Sellar, S. & Lingard, B. (2013a). Looking East: Shanghai, PISA 2009 and the reconstitution of reference societies in the global education policy field. *Comparative Education, 49*(4), 464-485.
Sellar, S. & Lingard, B. (2013b). The OECD and global governance in education. *Journal of Education Policy, 28*(5), 710-725. doi: 10.1080/02680939.2013.779791
Sojourner, A. A., Mykerezi, E. M., & West, K. K. (2014). Teacher pay reform and productivity. *Journal Of Human Resources*, *49*(4), 945-981.
Springer, M. G. (Ed.). (2009). *Performance incentives: Their growing impact on American K-12 education*. Brookings Institution Press.
Springer, M. G., & Gardner, C. D. (2010). Teacher pay for performance context, status, and direction. *Phi Delta Kappan*, *91*(8), 8-15.
Statistical Yearbook. (2016). Retrieved from http://www.statssh.gov.cn/data/toTjnj.xhtml?y=2016e)
Storey, A. (2000). A leap of faith? Performance pay for teachers. *Journal of Education Policy, 15*(5), 509-523. doi: 10.1080/026809300750001667
Teacher Incentive Fund. (2015). Awards. Retrieved from http://www2.ed.gov/programs/teacherincentive/apps2012/index.html
TedGlobal. (2009, July). Dan Pink: The puzzle of motivation [Video file]. Retrieved from http://www.ted.com/talks/dan_pink_on_motivation
Temple, B. (2002). Crossed wires: Interpreters, translators, and bilingual workers in cross-language research. *Qualitative Health Research, 12,* 844-854.
Temple, B., & Edwards, R. (2002). Interpreters/translators and cross-language research: Reflexivity and border crossings. *International Journal of Qualitative Methods, 1*(2), 1-12. Retrieved from htpp://www.ualberta.ca/~ijqm
Temple, B., & Young, A. (2004). Qualitative research and translation dilemmas. *Qualitative Research, 4,* 161-178.
Temple B., Edwards R., & Alexander C. (2006). Grasping at context: Cross language qualitative

research as secondary qualitative data analysis. *Forum: Qualitative Social Research, 7*(4). Retrieved from http://www.qualitative-research.net/index.php/fqs/article/view/176/393

The New Teacher Project. (2014, July 15). *Shortchanged: The hidden costs of lockstep teacher pay.* Retrieved from http://tntp.org/publications/view/retention-and-school-culture/shortchanged-the-hidden-costs-of-lockstep-teacher-pay

Trice, H. M., & Beyer, J. M. (1993). *The cultures of work organizations.* Englewood Cliffs, New Jersey: Prentice-Hall.

Trujillo, T. (2014). The Modern Cult of Efficiency Intermediary Organizations and the New Scientific Management. *Educational Policy*, *28*(2), 207-232.

Tucker, M., (2011). *Surpassing Shanghai: An Agenda for American Education Built on the World's Leading Systems.* Cambridge: Harvard Education Press.

United States Department of Education. (2015). *Race to the Top Annual Performance Report.* Retrieved from http://www2.ed.gov/programs/racetothetop/performance.html

Vagi, R. (2014). Merit pay and music education: A motivation perspective. *Arts Education Policy Review, 115*(3), 98-103. doi:10.1080/10632913.2014.914395

Verger, A. (2014). Why do policy-makers adopt global education policies? Toward a research framework on the varying role of ideas in education reform. *Current Issues in Comparative Education 16*(2), 14-29.

Verger, A., Lubienski, C., & Steiner-Khamsi, G. (2016). The Emergence and Structuring of the Global Education Industry: Towards an Analytical Framework. In A. Verger, C. Lubienski, & G. Steiner-Khamsi (Eds.). *World Yearbook of Education 2016: The Global Education Industry*, 3-24. New York, NY: Routledge.

Vroom, V. H. (1964). Work and motivation. New York, NY: Wiley.

Walberg, H. J., & Bast, J. L. (2003). *Education and capitalism: How overcoming our fear of markets and economics can improve America's schools.* Stanford, CA: Hoover Institution Press.

Walker, P. (2013, January 15). Teachers' pay rises to be based on performance, Michael Gove confirms. *The guardian.* Retrieved from http://www.theguardian.com/education/2013/jan/15/teachers-pay-performance-michael-gove

Weiner, M. J. (1980). The effect of incentive and control over outcomes upon intrinsic motivation and performance. *Journal of Social Psychology, 112*(2), 247-254.

Xiu, L., & Gunderson, M. (2013). Gender earnings differences in china: Base pay, performance pay, and total pay. *Contemporary Economic Policy, 31*(1), 235-254. doi:10.1111/j.1465-7287.2011.00307.x

Xu, L. C. (2000). Control, incentives and competition. *Economics of Transition, 8*(1), 151-173.

Yarovaya, Y. B. (2015). Standardization of Primary Education in Great Britain. *European Journal of Contemporary Education, 12*(2), 169-174. doi: 10.13187/ejced.2015.12.169

Yin, R. K. (2003). *Case Study Research: Design and Methods.* Thousand Oaks, CA: Sage Publications.

APPENDIX A: IRB LETTER

UNIVERSITY OF ILLINOIS
AT URBANA-CHAMPAIGN

Office of the Vice Chancellor for Research

Office for the Protection of Research Subjects
528 East Green Street
Suite 203
Champaign, IL 61820

March 10, 2016

Christopher Lubienski
Department of Education Policy, Organization and Leadership
338 Education Bldg
1310 South Sixth Street
Champaign, IL 61820

RE: *How Performance Incentive Shape Culture in High Performing Schools*
IRB Protocol Number: 16629

Dear Dr. Lubienski:

This letter authorizes the use of human subjects in your project entitled *How Performance Incentive Shape Culture in High Performing Schools*. The University of Illinois at Urbana-Champaign Institutional Review Board (IRB) approved, by expedited review, the protocol as described in your IRB application. The expiration date for this protocol, IRB number 16629, is 03/07/2017. The risk designation applied to your project is *no more than minimal risk*.

Copies of the attached date-stamped consent form(s) must be used in obtaining informed consent. If there is a need to revise or alter the consent form(s), please submit the revised form(s) for IRB review, approval, and date-stamping prior to use.

Under applicable regulations, no changes to procedures involving human subjects may be made without prior IRB review and approval. The regulations also require that you promptly notify the IRB of any problems involving human subjects, including unanticipated side effects, adverse reactions, and any injuries or complications that arise during the project.

If you have any questions about the IRB process, or if you need assistance at any time, please feel free to contact me at the OPRS office, or visit our Web site at http://oprs.research.illinois.edu.

Sincerely,

Rebecca Van Tine

Rebecca Van Tine, MS
Human Subjects Research Specialist, Office for the Protection of Research Subjects

Attachment(s): Written informed consent document, and Waiver or Alteration of Informed Consent form

c: Priya La Londe

APPENDIX B: INTERVIEW PROTOCOL

Purposes of interview/访谈的目的

The purpose of the interview is to better understand how teachers, principals, and other administrators make sense of performance incentives. Specifically:

1. How do participants understand the aims of performance incentives? (Q1, Q2)
2. What are participants' experiences with performance incentives? (Q2, Q3)
3. In what ways do performance incentives contribute to teaching/leadership effectiveness for this teacher/administrator? (Q4)
4. What are the participants' experiences with relational trust and teamwork among their colleagues? (Q5)
5. In what ways do performance incentives shape trust and teamwork? (Q6)

访谈的目的是为了更好地了解教师、校长、和其他管理者如何理解绩效奖励（绩效工资）。具体如下：

1、参与者（教师、校长、其他管理者）如何理解绩效奖励的目的?(Q1、Q2)

2、参与者（教师、校长、其他管理者）对绩效奖励有怎样的体会?(Q2、Q3)

3、绩效奖励是通过什么样的方式使教师和其他管理者提高教学领导效能？（Q4）

4 、关于同事之间的信任关系、团队工作，参与者有怎样的体验？（Q5）

5、绩效奖励（绩效工资）在以怎样的方式形成同事之间的信任和团队工作？（Q6）

Introductions, informed consent, permission to tape, Background/介绍、知情同意、录音许可

A. [Ice-breaker, could also use some other comfortable question to get started]. Before I started my doctorate work, I also worked in schools. I always enjoy chatting with educators. It is my pleasure to be here in Shanghai and learn about you. Could you share with me how you decided to become an educator?

A.（"破冰船"，也可以使用其他一些舒适的问题开始）。在我读博士之前,我在学校工作过，我总是喜欢和教育工作者聊天。我很高兴能够在上海认识您。您能够和我分享一下您为何愿意成为一名教育工作者呢？

B. I would like to learn more about your experiences in education. Can you tell me more about your experiences? How long have you taught? Which ages? Which schools? How long have you been a school leader? Which ages? Which schools? Where did you receive your training?

B、我想了解更多关于您的教育经历。您能告诉我更多关于您的经历吗?您任教多长时间

了？从多少岁开始？您所在的学校是？您成为这个学校的领导者有多长时间？从几岁开始？您所在的学校是？在此之前您在哪里接受培养或培训过？

Aims/purposes of performance incentives

Now I would like to talk about performance incentives. Performance incentives are financial bonuses given to teachers and school leaders whose students perform well. This could be a bonus for an individual teacher, a group of teachers, or all the teachers and leaders in an entire school. Said another way, these bonuses are incentives given to teachers for improved or superior student performance. I am interested in learning about your perspectives on and experiences with performance incentives. So let's begin, ok?

1. How would you describe the purpose of performance incentives?
    a. How did you learn about performance incentives? Did the Ministry of Education or your principal explain incentives to you?
    b. When did you first learn about performance incentives?
    c. Please describe the performance incentive policy in your school as you understand it.
    d. What do you understand as the process for receiving a bonus?
    e. What other policies reward teachers and leaders for good student performance?

2. What is your overall view of performance incentives?
    a. What do you think are some benefits of awarding incentives to teachers and school leaders for student performance?
    b. How would you describe the challenges associated with awarding incentives to teachers and school leaders for student performance?

Brief member check on questions 1 & 2

绩效奖励的目的

现在，我想谈论绩效奖励（绩效工资）。绩效奖励（绩效工资）就是给予学生表现出色的教师和学校领导一定经济上的奖励。这种奖励可以是给予教师个体、教师群体、或者整个学校的所有老师和领导。换句话说，这些奖金是奖励给那些学生表现突出或者学生有所进步的教师们的。因此，我非常想了解您对于绩效奖励（绩效工资）有哪些个人的想法和体验。我们现在开始访谈，可以吗？

1、您认为绩效奖励有怎样的目的？

a. 您之前怎样理解绩效奖励？教育部门（教育部）或者你们的校长向你们解释过绩效奖励（绩效工资）这一措施吗？

b. 您第一次知道（或者了解）绩效奖励（绩效工资）是在什么时候？

c. 您是否可以描述一下就您所理解的，学校目前实施的绩效奖励（绩效工资）政策

d. 您所理解到的获取奖金的过程是什么样的？

e. 对于学生表现好的教师和领导，还有没有其他的奖励政策（或者措施）？

2、您对绩效奖励（绩效工资）的总体看法是什么？

a. 您认为给学生表现好的教师和领导授予奖励，会带来哪些好处？

b. 您认为对于学生表现好的教师和领导予以奖励的同时，会带来哪些挑战？

再快速检查一下问题1和2 。

Experiences with performance incentives

3. What are your personal experiences with performance incentives?
    a. Please describe the incentives you have received.
    b. How much was the bonus?
    c. When did you receive the bonus?
    d. How does this bonus shape your overall salary?
    e. What do you understand to be the reason you received this bonus? If student performance, what specific performance warranted this bonus?
    f. Who made this bonus allocation decision?
    g. In your opinion, how is your experience with bonuses similar to or different from your colleagues' experiences with bonuses? Please explain.

4. What are your perceptions of how incentives contribute to teaching and leadership effectiveness?
    a. How do incentives shape your teaching? Approach to teaching? Commitment to teaching? Motivation to learn about teaching?
    b. How would you evaluate the value or contribution of incentives to your overall teaching effectiveness?
    c. What else would help you be the best teacher you can be?

Brief member check on questions 3 & 4

对绩效奖励的体验

3、对于绩效奖励您有什么个人的体验？

a. 请您谈谈您得到过什么样的绩效奖励。

b. 奖金多少？

c. 您什么时候得到这笔奖金？

d. 这笔奖励如何改变您的整个工资？

e. 您理解您得到这笔奖金的缘由是什么（您知道您是因为什么得到这笔奖金的吗）？如果是因为学生，您认为学生表现如何才能让您得到奖金？

f. 谁来决定奖金的分配？

g. 在您看来，当获得奖金时，您的体验和同事的体验有何异同？？请详细阐述下。

4、你如何理解绩效奖励有助于提高教学与领导的有效性？

a. 绩效奖励是如何影响（或者塑造）您的教学？如教学的方式方法？对教学的认同感？学会教学的愿望？

b. 对于您整体的教学效能，您如何评价绩效奖励的价值与贡献？

c. 还有什么能够帮助您成为一名好老师？

快速检查问题3和4。

How performance incentives shape trust and teamwork

In the last part of this interview I want to talk about how incentives shape the trust between you and your colleagues and how incentives shape your teamwork. A lot of research tells us about what incentives are and the goals of incentives. But we don't know very much about how incentives shape how teachers collaborate together and approach each other? So let's begin, ok?

5. Relational Trust and Teamwork
    a. How would you describe your relationships with your fellow teachers? Principal? Other school leaders?
    b. In what ways does trust matter for in your work?
    c. Can you give an example of 1-2 colleagues whom you trust a great deal?
    d. In what ways do you work with your colleagues?
    e. In your view, how does this teamwork contribute to your teaching effectiveness?
6. Role of performance incentives in relational trust and teamwork
    a. In your view, how do financial bonuses shape your relationships with your colleagues? With your principal? Can you give an example of this?
    b. How do bonuses for student performance shape the work you and your colleagues do together?
    c. Think about a time when you received a bonus. How did your colleagues learn about your bonus? What made you decide to take this approach? Or if the Ministry or your principal shared this information, in what ways did your colleagues react?
    d. Think about a time when you learned one of your colleagues received a bonus. Can you describe how you learned about this? Can you describe your reactions?
    e. What other issues and concerns do you have related to performance incentives and your work or your work with your colleagues?

Brief member check on questions 5 & 6

绩效奖励如何塑造（或者影响）信任与团队合作？

在访谈的最后一部分，我想了解绩效奖励（绩效工资）如何影响（或者塑造）您和同事之

间的信任关系，如何影响您的团队工作。很多研究是阐述绩效工资的内涵和绩效工资的目标。但是，关于绩效工资如何推动教师间的合作和密切教师间的相互关系，我们知之甚少。我们现在开始访谈，可以吗？

5、关系信任和团队合作

a. 您觉得您和各位老师的关系如何?和校长的关系如何？和其他学校领导的关系如何？

b. 信任是以什么方式影响您的工作的？

c. 说说您非常信任的1-2同事。

d. 您以什么方式和您的同事一起工作？

e. 在您看来，团队合作怎样有助于您的教学效能的提高？

6、在关系信任和团队合作中，绩效奖励的作用

a. 在您看来，经济上的奖励怎样影响你和同事之间的关系，和校长之间的关系？请举一例

b. 因学生的表现而获得的奖金如何影响你和同事之间的团队合作？

c. 当你得到奖金时，你的同事是如何获知这一信息的？如果是您告知同事的，是什么促使您告知同事的，是什么促使您做出这一决定的？或者如果教育部或者校长分享（或者公布）这一信息时，你的同事反应如何？（你的同事以什么样的方式反应？）

d. 想想当您知道您的一个同事得到奖金时，您是如何获知的？您能描述您当时的反应吗？

e. 关于绩效奖励以及你和同事之间的团队合作，您还有什么其他的问题吗？

快速检查问题5 & 6

APPENDIX C: KEY TERMS AND RECRUITMENT QUESTIONNAIRE

Table B1
Key Terms Translation

English Term	English Definition*	Mandarin Pin Yin	Mandarin Characters
salary	overall composite of compensation a teacher received in a fixed salary distribution period	gong zi	工资
base pay	the total fixed salary that a teacher receives in a fixed salary distribution period	ji chu xing gongzi	基础性工资
merit pay	the total variable salary that a teacher receives in a fixed salary distribution period	ji xiao gong zi	绩效工资
merit pay bonus	the variable salary determined by the school administration that a teacher will receive in a fixed salary distribution period	jiang li xing ji xiao gongzi	奖励性绩效工资
district bonus	the variable salary determined by the district leadership that a teacher is eligible to receive in at the end of the first and second semester	qu jiang li	区奖励
city bonus	the variable salary determined by the city leadership that a teacher is eligible to receive one time at the end of the school year	shi jiang li	市奖励
motivation	desire, willingness, or reasons to do something	dong ji	动机
fair	what is just, appropriate, legitimate, or within the rules and standards	gong ping	公平
love	deep affection, interest, or pleasure	ai xin	爱心
enthusiasm	eager enjoyment or excitement	re qing	热情
passion	intense desire or enthusiasm	ji qing	激情
morality	principles held by a person, group or a society, concerning the distinction between right and wrong or good and bad behavior	dao de	道德
conscientiousness, conscientious	the act of doing what is right or to do one's work well or thoroughly	liang xin	良心
responsibility	a duty to fulfill the obligations and assignments prescribed to a teacher	zeren	责任
teaching quality	a teacher's skills and aptitude	jiao xue zhi liang	教学质量
team work	two or more teachers working together	tuan dui he zuo	团队合作

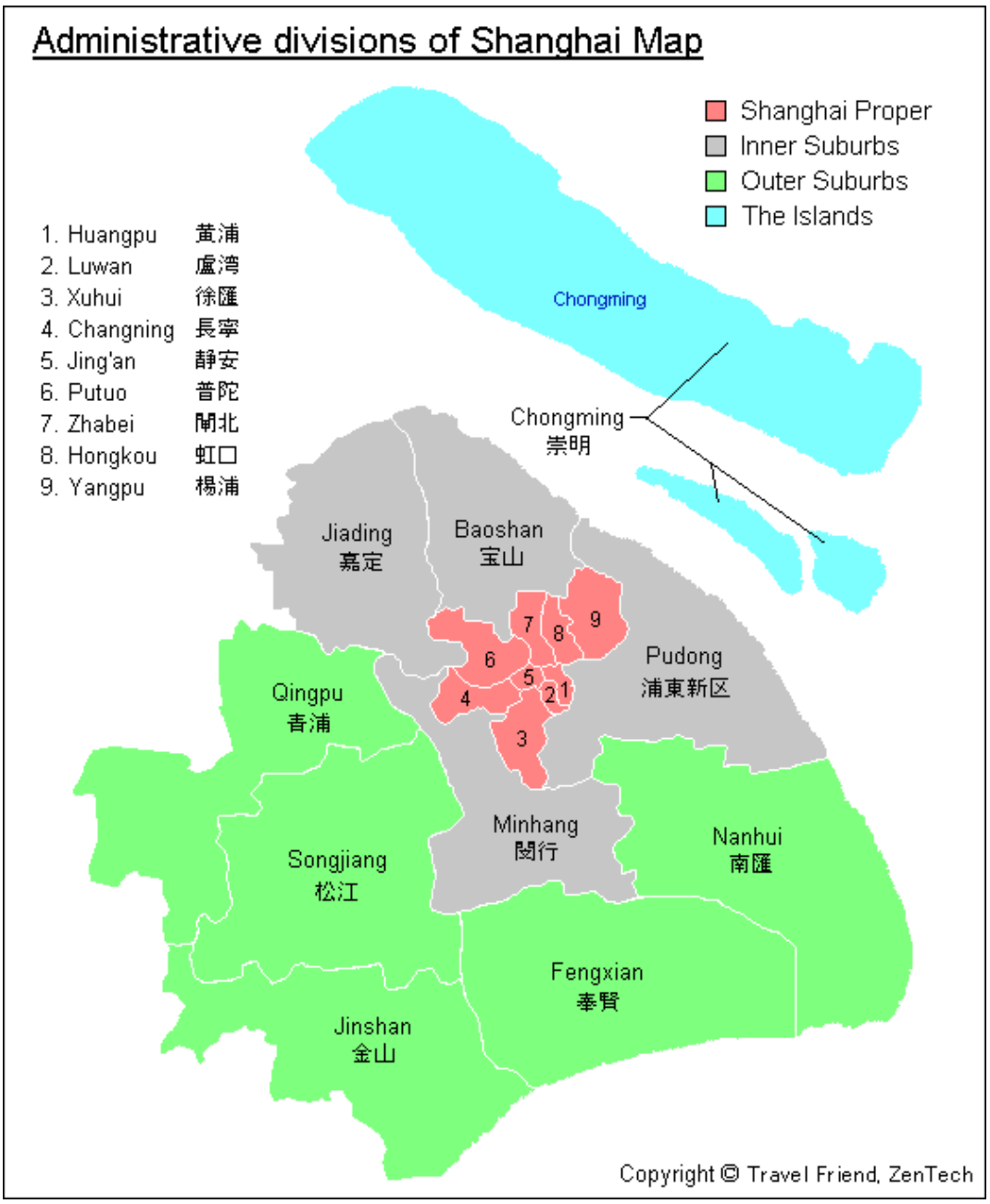
relationship	the interactions, relations, and communications between two or more people	guan xi	关系
working hours	the number of class periods a teacher teaches in one week	gong zuo shi chang	工作时长
number of classes	the number of cohorts of students a teacher is responsible for teaching in one academic year	ke shi liang	课时量
competition	events that occur throughout the school year wherein students and teachers from schools in neighboring districts or provences compete against one another in mathematics, science, Chinese, English, art, music, and physical education	bi sai jing sai	比赛／竞赛
teacher competition	events that occur throughout the school year wherein teachers from schools in neighboring districts or provences compete against one another to showcase pedagogy and instruction in their subject area expertise	jiaoxuejingsaishi fan ke gong kai ke	教学竞赛／示范课／公开课
student competition	events that occur throughout the school year wherein students from schools in neighboring districts or provences compete against one another in mathematics, science, Chinese, English, art, music, and physical education	ti yu bi sai ke ji jing sai wenti lei bi sai	体育比赛／科技竞赛／文体类比赛
veteran teacher	a teacher with five or more years of teaching experience	lao jiao shi	老教师
novice teacher	a teacher with four or less years of teaching experience	xin jiao shi	新教师
purpose	the aims of a particular policy, program, or act	mu di	目的
Head Teacher	the teacher in charge of general matters of a class (section) of students, who usually also teaches one subject to the students of that class	ban zhu ren	班主任
Pioneer battalion counselor	the counselor for a school's Young Pioneers Organization	da dui fu dao yuan	大队辅导员
Young Pioneer / Red Scarf Organization	a mass youth organization for children aged six to fourteen under the Communist Party of China; usually every elementary student becomes a member sooner or later	shao xian dui hong ling jin	少先队 / 红领巾

Note. * These terms were adapted from their broader meanings and applied to the context of teachers and teaching.

Table B2.
Recruitment Questionnaire and Information Sheet

English	Chinese Mandarin
My name is Priya La Londe. I am a doctoral candidate at the University of Illinois at Urbana-Champaign, a large research university in the United States of America. I am conducting research on how financial bonuses and other incentives shape school culture in high performing schools in Shanghai, China. This study is important because education policy around the world is moving toward awarding bonuses or merit pay to teachers and administrators whose students perform well on standardized assessments. I have two goals for this research. First, I want to describe the scope and aims of bonuses/merit pay and how they are used in schools. Second, I want to analyze how how bonuses/merit pay shapes school culture – specifically in terms of trust and collaboration among teachers. I will not collect any identifiable data on students or teachers. I am not here to judge or evaluate your school or you work. I simply want to better understand your experiences with merit pay and your thoughts about how these bonuses shape your work.	我叫Priya La Londe，是一名来自伊利诺伊大学厄巴纳——香槟分校的博士研究生。伊利诺伊大学是美国一流研究型大学。目前我正在研究财政补贴和其他激励措施如何塑造校园文化，研究范围是中国上海的优质学校。本研究的重要性在于，目前世界各地的教育系统都在致力于推行以学生表现为评价标准的奖金制度。本研究主要有以下两个目的：首先，我希望了解贵校设立奖金的目的，奖金的内部结构以及奖金制度的运作情况。其次，我想要分析在这个过程中奖金制度如何塑造校园文化——特别是如何影响教师之间的信任关系与合作关系。我只是一名研究者，并非想要通过此研究来判断或者评价您的学校以及您的工作，过程中我也不会收集任何教师或者学生的私人信息。我只是希望能够了解您的经历、您对奖金的看法以及奖金如何影响您的工作。
Your Name, Your Mobile Phone Number, Your Email ID	您的名字 , 您的电话号码您的, 箱 , 微信
1) Have you ever received merit pay? If yes, when? How much merit pay?	(1)您曾经获得过奖金吗？若有，是什么时候获得？金额是多少？
2) In 2014 or 2015, did any teachers in your school received merit pay? If yes, approximately how many teachers? If yes, approximately how much merit pay did each teacher receive?	(2)在2014年和2015年期间，贵校有教师获得奖金吗？若有，大概有几个人获得？每人大约多少金额？
3) Do you consider your school a high ranking or top performing school in Shanghai? If yes, how do you know?	(3)在您看来，贵校在上海算是排名靠前或者说是优质学校吗？如果是，您是怎么知道的？
4) Would you be willing to participate in an interview with the Research Assistants and I about your experiences with merit pay? If yes, please suggest the most convenient date and time for you.	(4)您是否有兴趣参加访谈，向我和我的研究助手分享您关于奖金制度的体验和经历。如果您愿意，可以告诉我们您最方便的时间。

APPENDIX D: MAP OF SHANGHAI ADMINISTRATIVE DIVISIONS



APPENDIX E: DESIGN FEATURES OF PERFORMANCE-BASED COMPENSATION PROGRAMS IN THE U.S.

Program Structure					Performance Measures	Pay Allocation
Region and Program	**Starting Year**	**Budget in first year**	**Impact**	**T = Tournament F = Fixed Performance Contract**	**I = Inputs ` O = Outputs E = Supervisor Evaluation**	**Bonus Amounts**
States						
Florida	2007	$147.5 million	54 out of 67 districts	F	O, E	5-10 percent of average teacher annual salary
Minnesota	2004	$64 million	1/3 of students in the state	F	I, O	
New Mexico Districts	2004		All public schools	T, F	I, O, E	
Austin	2005	$9.7 million	Any public school where 75% of staff consents	F	I, O	$1500 per Student Learning Objective achieved per teacher and $2000 per subject per teacher for reading and math
Denver	1999	$25 million	Voluntary for all teachers	F	I, O	
Houston					O	

Note. Adapted from "Design and Implementation Considerations for Alternative Teacher Compensation Programs," by A. Brodsky, D. DeCesare, & J. Kramer-Wine, J., 2010, *Theory Into Practice, 49,* p. 213–222.

APPENDIX F: EXPERIMENTAL EVIDENCE ON PERFORMANCE-BASED COMPENSATION

PBC Program	Incentive Level	Citation	Method	Program Timeline	Findings
New York City, New York, USA	Schoolwide	Goodman and Turner, 2011	RCT	2007-2010	There was no effect on student achievement among students in the treatment group. Teacher absenteeism dropped slightly among those in the treatment group who received the largest incentives.
New York City, New York, USA	Schoolwide	Fryer, 2011	RCT	2007-2010	There was a negative but insignificant effect on student achievement, attendance, and graduation among students in the treatment group. There was a negative but insignificant effect on achievement among students in the treatment group in in larger schools. There was a negative but insignificant effect on teacher behaviors among teachers in the treatment group.
New York City, New York, USA	Schoolwide	Marsh et al., 2011	RCT	2007-2010	There was a negative but insignificant effect on student achievement, attendance, and graduation among students in the treatment group. There was a negative but insignificant effect on achievement among students in the treatment group in in larger schools. There was a negative but insignificant effect on teacher behaviors among teachers in the treatment group.
Nashville, Tennessee, USA	Individual math teachers	Springer et al., 2010	RCT	2006-2009	Student in the treatment group did not significantly outperform students in the control group. Teachers in the treatment group were no more likely to report that incentives discouraged teachers from working together.
Andhra Pradesh, India	Individual and Schoolwide	Duflo, Hanna, and Ryan 2012	RCT	2005-2010	There was a positive effect on teacher attendance and student achievement among teachers and students in the treatment group.
Andhra Pradesh, India	Individual and Schoolwide	Muralidaran and Sundararaman, 2011	RCT	2005-2010	Students in the treatment group had higher scores in non-incentive subjects – science and social studies. Students in the treatment group whose teachers received schoolwide incentives did not significantly improve in math but improved in language.
Kenya	Schoolwide	Glewwe et al.,	RCT	1998-1999	Students in the treatment group improved performance, but this

		2003			was not sustained the year after the program ended. Attendance improved among teachers in the treatment group slightly.
Israel	Schoolwide	Lavy, 2002	Regression discontinuity	1996-1997	Students in the treatment group showed reduced rates in dropout and increases in the number of credits taken.
Israel	Individual	Lavy, 2009	Regression discontinuity	2000-2001	There was a slight positive effect on student achievement in math and English among students in the treatment group. Teachers in treatment groups were more likely to work extra time with low performing students, use small group instruction, and differentiated instruction.
Little Rock, Arkansas, USA	Individual	Winters et al., 2007	Difference-in-difference	2003-2006	Students slightly improved their math performance.
England, U.K.	Individual	Atkinson et al., 2009	Difference-in-difference	1999-2002	Students whose teachers had more experience showed higher test scores.
Michigan, USA	Individual	Eberts et al., 2002	Difference-in-difference	1996-1999	Student improvement in course completion was the direct result of PBC. Student decline in attendance was not the direct result of PBC.
North Carolina, USA	Schoolwide	Vigdor, 2008	Regression discontinuity	1996-2007	Students whose teachers did not receive a bonus improved slightly.
Dallas, Texas, USA	Schoolwide	Ladd, 1999	Difference-in-difference	1991-1995	Students improved in dropout, attendance, and slightly in math and reading. Principal turnover rates increased.
USA	Individual	Figlio and Kenny, 2007	OLS regressi	1988 NELS and	There was no effect on student achievement among students in schools with group- or school-based bonuses. There was a

			on	1993 SAS	slightly significant slightly positive effect on student achievement among students in schools with individual bonuses. The largest positive effects on student achievement were in schools with individual bonuses and with the most students from low-income families.
28 OECD nations	Schoolwide and individual	Woessman (2011)	OLS regression	2003	Students whose teachers received merit pay performed slightly better.

Note. Adapted from "The Economics and Politics of Teacher Merit Pay," by A. Leigh, 2012, *CESIfo Economic Studies, 59*, p. 1-33.

APPENDIX G: MINHANG DISTRIC 2014-2015 SCHOOL YEAR RANKINGS

2014学年闵行区学校、幼儿园办学绩效评价结果公示

发布时间：2015-11-11　来源：

根据《2014学年闵行区学校、幼儿园办学绩效评价方案》要求，经局绩效评价工作小组评价、领导小组审定，现将办学绩效评价结果予以公示。 公示时间：2015年11月11日—11月16日。

2015年11月11日

2014学年闵行区学校、幼儿园办学绩效评价奖励名单

一、综合奖（共106所）

奖 项	学 段		学 校
优秀一等（26所）	幼儿园	9所	景谷二幼、佳佳中心幼、莘庄幼儿园、虹鹿（虹桥）幼儿园 龙茗路幼园、七宝中心幼、吴泾一幼、古美中心幼 龙柏二幼
	小学	8所	田园外小、实验小学、平南小学、莘庄镇小、明强小学、江川小学华坪小学、汽轮小学
	初中	5所	梅陇中学、实验西校、基地附中、莘松中学、文来中学（初中部）
	九年一贯	2所	莘城学校、莘光学校
	高中阶段	2所	七宝中学、群益职校
优秀二等（30所）	幼儿园	10所	闵行四幼、闵行一幼、鑫都幼儿园、康城幼儿园、君莲幼儿园 华漕中心幼、莘松幼儿园、天恒名都幼、纪王幼儿园 华漕金色幼
	Elementary school	7所	交大实小、蔷薇小学、新梅小学、闵行小学、莘松小学、鹤北小学浦江一小
	初中	6所	七宝三中、七宝实中、上虹中学、北桥中学、七宝二中、上宝中学
	九年一贯	4所	明星学校、康城学校、启智学校、启音学校
	高中阶段	3所	闵行中学、莘庄中学、成教一中心
优秀三等（39所）	幼儿园	13所	颛桥幼儿园、安宁路幼园、万源城幼儿园、浦江三幼 鹤庆幼儿园、莘庄二幼、星辰幼儿园、梅陇镇中心

			七宝实验幼儿园、春欣幼儿园、龙柏一幼、浦江一幼 平吉四街坊幼儿园
	Elementary school	**11所**	**平阳小学、花园学校、虹桥小学、明强二小、梅陇小学、颛桥小学罗阳小学、七宝实小、紫竹小学、航华二小、七宝外小**
	初中	7所	华二初中、闵行四中、龙茗中学、龙柏中学、浦江一中、鹤北初中文绮中学（初中部）
	九年一贯	4所	君莲学校、纪王学校、古美学校、少体校
	高中阶段	4所	闵行二中、田园高中、闵行三中（高中部）、文来中学（高中部）
进步奖 （11所）	**Elementary school**	4所	吴泾三小、碧江小学、丽江小学、浦航小学
	初中	3所	友爱中学、航华二中、罗阳中学
	九年一贯	2所	强恕学校、华漕学校
	高中学校	2所	浦江高中、金汇高中

二、 创新创效奖和特殊贡献奖（共3项）

类别	项目/成果	
创新创效奖	1	平南小学：数字化教材开发
	2	七宝中心幼儿园：家校互动平台建设
特殊贡献奖	1	青少年活动中心：学生校外活动品牌打造

【备注】民办中小学列入了本次评价范围， 评价结果一并公布，但不发奖励经

APPENDIX H: SCHOOL M1 PARTICIPANTS INVENTORY

Participant	title	grade level(s)	subject	lessons	leader	gender	years experience	years at M1	hometown
1	administrator		n/a	0		female	25	4	Shanghai
2	teacher	3	Chinese	13	bzr	female	23	8	Shanghai
3	teacher	3	English	14	rd	female	3	3	Shanghai
4	teacher	1, 2	science		grs	female	2	2.5	Shanghai
5	teacher	3	physical education	11	grs	female	12	4	Shanghai
6	teacher	4	Chinese	12	bzr	female	32	10	Shanghai
7	teacher	5	English	10	ss	female	12	8	Shanghai
8	teacher	3	information technology	11		male	4	4	Shanghai
9	teacher	2	science	14	ad	female	10	10	Shanghai
10	teacher	3	math	13	grs	female	12	8	Jiangsu
11	teacher	2	Chinese	13	bzr; grs	female	6	4	Heilongjiang
12	teacher	1, 2	Music	16	ss	female	16	3	Jiang xi
13	teacher administrator	1, 2	Art	17	ss	female	22	10	Shanghai
14	teacher	3	English	13	grl	female	2	2	Shanghai
15	teacher	4	math	15	bzr	female	1	1	Shanghai
16	teacher	1	Chinese	16	ss	female	15	15	Shanghai
17	teacher	5	art	12		female	8	6	Shanghai
18	teacher	4	music	11		female	7	5	Shanghai
19	teacher	5	Chinese	13		female	6	6	Shanghai
20	teacher	1	math	14	bzr	female	3	3	Shanghai
21	teacher	2	art	12		female	4	4	Shanghai