


# Why Merit Pay Fails: Comparison of Public and Private School Practices in Shenzhen, China

ECNU Review of Education  
2024, Vol. 7(3) 809–832  
© The Author(s) 2022  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/20965311221144461  
journals.sagepub.com/home/roe



Hejia Shi (史和佳) 

The University of Hong Kong

Dan Wang (王丹)

The University of Hong Kong

## Abstract

**Purpose:** Accounts for failures of merit pay in educational settings were stagnated by the “political obstruction hypothesis,” which blamed teacher unions for impeding the merit pay schemes from proper function. It required stronger evidence from both public and private schools to refute this hypothesis.

**Design/Approach/Methods:** Accountability typology was conceptualized to uncover the power relations underpinning merit pay schemes, and teachers’ motivational status was described from the perspective of Self-Determination Theory (SDT) to analyze the corresponding consequences. One public school and three private schools were purposefully sampled, and data included principal interviews, teacher interviews, school documents, and relevant policy documents.

**Findings:** Our findings suggested an opposite conclusion to the “political obstruction hypothesis”

---

The research was done and the manuscript was submitted when the corresponding author was studying as a Ph.D. candidate at the University of Hong Kong. Corresponding address below is updated to reflect where the author is currently working and living.

---

### Corresponding author:

Hejia Shi, Shifu Avenue 1139, Jiaojiang District, Taizhou, Zhejiang 318000, China.

Email: bnushihj@qq.com



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access page (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

that the more dependent on market, the more intense the conflict between external accountabilities and the nature of teaching. Thus, we supported the “nature of teaching hypothesis” that merit pay was inapplicable in educational settings due to the inherent contradictions between market accountability and professional accountability.

**Originality/Value:** Policymakers and school administrators had been misled by the neoliberal initiative of borrowing merit pay into educational settings to promote the quality of teaching. Our findings provided strong evidence to refute this hypothesis and to restore the appreciation of professional autonomy.

### Keywords

Market influence, merit pay, teacher accountability, teacher motivation

Date received: 14 August 2021; revised: 24 January 2022; accepted: 8 April 2022

## Introduction

Merit pay as a compensation increase based on measurement of one’s “performance” has been adopted in public administrations in 80% of the Organisation for Economic Co-operation and Development (OECD) countries (Dielmann, 2012; Lah & Perry, 2008; Ruffini et al., 2020). China also gradually embraced merit pay in educational policies to break the egalitarianism in teachers’ remuneration along with the economic reform and opening-up. In 1985, the “structural salary” schedule was rolled out by the government to connect teachers’ remuneration with differed actual workload. In 1993, a professional career ladder was introduced in the “Hierarchical Salary of Professional Technical Post” policy linking teacher salary with professional title appraisal in 1993 (He, 2019). A deepening merit pay reform in compulsory education was later launched in 2009 to compensate teachers based on the results of the performance appraisal and motivate them with pecuniary incentives (Ministry of Education of China [MOE], 2008). These reforms were expected to improve teaching practices, encourage cooperation, and attract talents to the profession (Fan & Fu, 2011).

Since 2009, however, researchers have found relentless complaints among teachers about merit pay. It was criticized for unfair assessment criteria for non-exam subjects (e.g., music, PE), preferential biases toward administrative positions, and the overemphasis on students’ academic performance (X. Li, 2010; K. Li, 2013; Liu, 2012; Liu et al., 2016; Wang et al., 2014; Xiong, 2009; Zhang, 2010). Meanwhile, evidence about its impacts on teacher development was at best inconclusive, if not negative. While some case studies found teacher development possible under merit pay reforms (e.g., Wang, 2007; Zhang & Ng, 2017), the success could hardly be attributed to the pecuniary incentives or the summative performance appraisal, but rather the promotion of collaborative

culture or the developmental use of appraisal results. More frequently, researchers found negative impacts of performance-contingent incentives, including worsened test-oriented climate, wide use of the appraisal results for punitive instead of developmental purposes, overreliance on quantifiable indicators at the cost of a qualitative appraisal, and accentuation on individual incentives while overlooking collective stimulation (Fan & Fu, 2011; Meng & Yuan, 2014). The Chinese literature echoes international scholarship, portraying merit pay as a policy failure (Ballou & Podgursky, 1993; Gius, 2013; Goldhaber et al., 2008; Yuan et al., 2013).

There are roughly two schools explaining the failure of merit pay in public schools: the “nature of teaching” hypothesis and the “political obstruction” hypothesis (Goldhaber et al., 2008). The different explanations vary significantly in their practical implicatures: The “nature of teaching” hypothesis rejects merit pay in all educational settings, while the “political obstruction” hypothesis defends merit pay even in public schools with failed examples. However, the extant literature lacks further evidence to bridge these two contesting arguments, leaving the debate stagnated with wavering policy directions. This study attempted to bridge the empirical gap by comparing the practices of merit pay in public and private schools with the analytical framework of accountability typology, trying to figure out what the most essential reason for the failures of teachers’ merit pay is, and how it manifests itself in terms of multiple accountabilities.

### **Accounts for failures of merit pay**

Merit pay may be dysfunction when it lacks a democratic foundation or measurement accuracy. For instance, employees may become discouraged, less productive, or even querulous when they perceive the performance measured by the manager as underrated due to subjective factors (Dessler, 2011), or when individuals are held accountable for performance that depends more on teamwork or other external factors that are beyond one’s control (Noe et al., 2011). Yet, these are at most critiques of technical failures that simply call for improvement to the process and instruments.

More essential defect of merit pay lies in its narrow focus on observable behavioral index over unobservable aspects due to its methodological limits. Even when merit pay schemes are implemented as intended to orient employees’ behavior toward prescribed objectives, the narrow focus on observable aspects over unobservable aspects may still decline the overall quality of work, especially when the unobservable aspects are essential to the success of the task (Al-Ubaydli & Lee, 2012). The emphasis on accountability and incentives may also discourage long-term professional development while attaching stakes to the emergent performance goals (Gailmard & Patty, 2007; Hood, 1991).

In addition to the methodological defect, merit pay could also be problematic in public sectors due to the motivational misalignment. Since employees in public sectors are often motivated by

intrinsic incentives rather than extrinsic incentives (Cacioppe & Mock, 1984; Crewson, 1997; Frey & Osterloh, 2013), the introduction of and emphasis on behavioral objects and pecuniary incentives may work as “substitutes” rather than “complements” for employees’ pro-social motivations when the treatment indicates “bad news” about the manager, triggers “moral disengagement,” or evokes “control aversion” in the targets, thus causing a “crowding-out” effect (Benabou & Tirole, 2003; Bowles & Polania-Reyes, 2012; Falk & Kosfeld, 2006; Frey, 1997; Frey & Osterloh, 2002; Lidenberg, 2001; Weibel et al., 2007, 2010).

Accounts for failures of merit pay in educational settings are also grounded in the aforementioned defects: the “political obstruction” hypothesis attributes the failure of merit pay to external obstruction from teacher unions, which prevents the schemes from proper functioning (Ballou, 2001), and the “nature of teaching” hypothesis attributes the failure of merit pay to the inherent methodological defects and the unintended “crowding-out” effect brought by motivational misalignment (Murnane & Cohen, 1986).

Researchers holding the political hypothesis argued that merit pay schemes were far more favorable among private school teachers than their public counterparts (Ballou & Podgursky, 1993). They interpreted such difference as that private schools differed from public schools in their management style (more flexible compensation system), political context (less engaged with teacher union), and dependence on market environment (more sensitive to market demands), rendering them in conditions closer to the private sector and more suitable for merit pay schemes (Ballou, 2001; Ballou & Podgursky, 1993). They tended to blame the technical failures on teachers’ obstructions, but avoid further examination of the methodological defects and motivational misalignment.

Researchers advocating the “nature of teaching” hypothesis insisted on teaching as a collective endeavor, depending on professional collaboration for sustained development (Paine & Fang, 2006; Paine & Ma, 1993). They perceived merit pay as flawed by methodological defects because student outcomes, the common-used performance indicator in merit pay appraisal, were the results of multiple interactive influences from family, school, and the larger social environment, far beyond the control of single teachers (Pham et al., 2021), and much of teachers’ daily effort that was critical for improving students’ well-being was invisible to school management and hard to be quantified (Hannaway, 1992; Warren & Ward, 2019).

Furthermore, although the crowding-in effect exists in theory when incentives align with teachers’ altruistic motivation and adhere to self-interested motivation, the crowding-out effect on intrinsic motivation was often more prominent, especially when teachers’ work was alienated under test-oriented or market-oriented climate (Bowles & Polania-Reyes, 2012; Weibel et al., 2010). Teachers’ preference may even change endogenously under the influence of merit pay to be lacking enthusiasm and recognition of educational values, and become more likely to work for better pay rather than the public good (Gius, 2013).

Teachers' merit pay reform in China was initiated in 2009, intending to attract quality new teachers and facilitate professional in-service development with pecuniary incentives. Yet, researchers have found that regional gap was even enlarged due to the localized fiscal sources for merit pay, and teachers were hardly motivated due to the problematic performance appraisal (Fu & Zeng, 2010; Tang & Hu, 2013). Existent domestic literature mainly attributed the failure of merit pay reform to the defects of fiscal system and distortion of policy design at school levels, lacking reflections on the fundamental hypothesis concerning the relation between pecuniary incentives and teacher motivation (Fu, 2013; Lai, 2013; Zhao & Liu, 2017). Empirical evidence on teacher supply has confirmed that influences of pecuniary incentives would be insignificant when teachers were satisfied with working conditions, implying the inadaptability of merit pay under certain circumstances (Huang et al., 2021). Yet, further studies are needed since current debate has not touched on fundamental hypothesis with empirical data.

### **Accountability typology of teachers' merit pay**

Essentially, merit pay holds teachers accountable for their "merit" or "performance." As a form of accountability institution, its values and policy goals vary by different types of accountabilities. Generally, researchers have identified three types of accountabilities relevant to teachers' merit pay: (1) administrative accountability, or described as organizational and bureaucratic accountability, that sets prescribed goals and procedures by top-down administrative power; (2) professional accountability that expects teachers to uphold professional norms and standards defined by the professional community; (3) market accountability that demands desired services including response to complaints and compliance with consumers' choice (Darling-Hammond, 2004; Williams & Engel, 2012).

Each type of accountability has correspondent indicators common in teachers' work: administrative accountability often inspects teachers' attendance, compliance, workload, and post-duty conducted by administrators; professional accountability embodies professional requirements via teaching contests, class observation, lesson study, and recognition of traditional qualifications as educational attainment, work experience (seniority), and academic publications; market accountability underpins the wide adoption of questionnaire to include parents and students as teacher appraisers, as well as various procedures to take care of parents' disputes with teachers or individualized appeals (Pang et al., 2019; Wang, 2005; Zhao et al., 2011).

The three accountabilities may influence teacher motivation in different orientations. Merit pay geared toward professional accountability has the potential to advance teacher professional development and encourage intrinsic motivation with incentives working as "complements." Yet, when teachers are pressured by merit pay plans orienting toward market and administrative purposes,

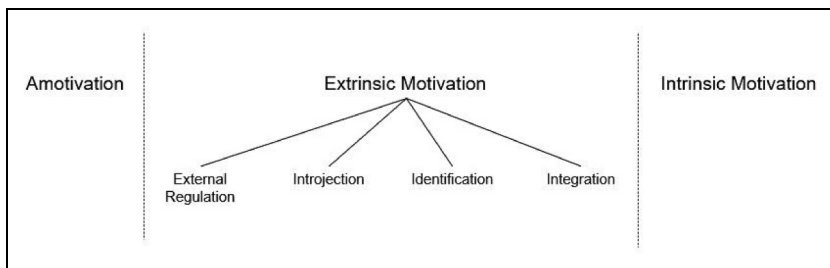
non-educational purposes may function as “substitutes” for pedagogical principles and result in crowding-out effect. Therefore, the accountability typology can be utilized to explain the varying consequences of different merit pay schemes by detailed examination of the relationship between merit pay structure and teachers’ agency.

## Self-determined theory and taxonomy of motivation

Although evidence has shown that students of teachers under merit pay are higher in test scores (Cooper & Cohn, 1997; Dee & Keys, 2004; Woessmann, 2011), it is inadequate to judge the effects of teachers’ merit pay since the quality of teaching is beyond mere cognitive development. Meta-analysis of merit pay schemes has found that there was little change in teachers’ instruction, number of working hours, job stress, or collegiality under merit pay schemes (Yuan et al., 2013). To better elaborate the delicate differences in teachers’ motivation, a taxonomy of motivation is borrowed from the self-determination theory (SDT; Ryan & Deci, 2000).

SDT distinguishes different types of motivation based on different reasons that drives action and argues that the motivational types matter in eliciting and sustaining people’s efforts (Deci & Ryan, 1985). Basically, motivation can be divided into two categories: intrinsic motivation which rewards people with inherent satisfactions in the task, and extrinsic motivation which relies on separable outcome to encourage efforts. Yet, in actual practice, people’s motivational status may shift on a spectrum of internalization where there are several transitional stages between the typical extrinsic and intrinsic motivation as shown in Figure 1 (Ryan & Deci, 2000).

The strength of motivation starts from “amotivation” as not motivated to “intrinsic motivation” as strongly and continually motivated, while “extrinsic motivation” contains several subcategories including “external regulation,” “introjection,” “identification,” and “integration.” The perceived locus of causality, namely the attributional source of drive, shifts from “external” at the “external regulation” stage to “internal” at the “integration” stage. For instance, a teacher at the “identification” stage may



**Figure 1.** The taxonomy of motivation.

Source. Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 61.

recognize the prescribed goals of test scores, which are extrinsic to teaching practices, because she believes the test-oriented strategy is helpful in maximizing students' chances into good colleges, which reflects somewhat internal locus of causality from her endorsement of goals.

It is noteworthy that teachers' motivation is likely to move either rightward or leftward as reactions to the dynamics in practice, so the impacts of merit pay on teachers' motivation can be shown either in the status or in the change of motivation. Empowered by the SDT, the analysis can trace the status and change of motivation back to different accountabilities and line up the data of public and private schools in a comparable framework.

### **Research setting and data**

So far, the debate between the "political obstruction" hypothesis and "nature of teaching" hypothesis is stagnated because the studies were conducted in either public schools or private schools separately and under different frameworks, lacking a bridging perspective to include comparable examples from both public and private schools (Ballou, 2001). In light of this gap, China offers a perfect scenario for this comparison. Its government has been promoting merit pay in public schools since 2009. The country features a strong public school system as well as a booming private educational sector. Public school teachers are members of the National Federation of Trade Unions without a profession-specific union. Therefore, teachers' union exists only nominally with minimal political power to influence teachers' work conditions (Gao et al., 2007; Xu & Yang, 2008). Embedded in substantial market influence and marginal union effect, China provides an ideal setting to test the two hypotheses under debate.

Teachers' remuneration in public schools in China is constituted by four parts since the merit pay reform: post-pay, level pay, merit pay, and subsidies, among which post-pay, level pay, and subsidies constitute the basic salary depending on teachers' professional title, while merit pay is meant to be the incentive salary depending on teachers' performance (Pang et al., 2019). Usually, performance appraisal plan that determines teachers' merit pay is deconcentrated to schools while local governments retain the power to intervene. Thus, school documents will provide detailed data about merit pay plan in public schools. On the contrary, merit pay in private schools is less subject to government policy with greater autonomy. The structure of teachers' remuneration in private schools may vary significantly depending on school context, using multiple synonyms for merit pay. Thus, investigation of merit pay in private schools shall be anchored in data and informed by the nature of merit pay rather than a specific pattern.

The study was carried out in Shenzhen, Guangdong Province, one of the most developed cities and a pioneer of education reforms in China. Since the exploratory interviews with principals suggested that teachers' remuneration is largely of identical structure among public schools while private school

teachers' remuneration is highly dependent on school management and school's orientation as for-profit or non-profit, four schools, one public, and three privates, were purposively sampled for comparison. School D was added as a case for the for-profit category since the principal of School C declined further interviews with teachers. Since High School Entrance Examinations (HSEE, *Zhong Kao* in Chinese) are no less competitive than College Entrance Examinations (CEE, *Gao Kao* in Chinese), and teachers in junior middle schools are exposed to similar test-oriented pressure as those in senior middle schools in general sense when contextual differences across schools are bracketed, the inclusion of School D will not weaken the comparability across case schools.

Data were collected through interviews with principals and teachers as well as from school policy documents. Since teachers of different subjects and different working experience may experience different pressures under merit pay, sampling of teacher participants purposefully included teachers of academic and non-academic subjects and different working generations in each case school. Table 1 reports the school sample and data forms, and Table 2 reports the background of teacher participants. All the interview participants were informed of their right to keep personal information confidential and withdraw from the research at any time. A written consent form was signed by interview participants before data collection, and a copy of interview transcript was proofread by interview participants afterward.

Data analysis adopted deductive interpretation, which is semi-structured by the accountability typology and the motivation taxonomy. The semi-structured analysis is composed of three key aspects: primary stakeholder, power relation, and practical impacts. The three aspects of analysis are neither clear-cut nor mutually exclusive but closely interlinked to portray the same object from different angles.

First, the analysis of primary beneficiary intends to figure out whose interest is mostly privileged behind a particular arrangement. It examines the purpose behind the selection of indicators, implementation of measuring activities, and application of results. In this way, merit pay schemes are dissected into relatively independent subsystems featured with a particular type of accountability.

**Table 1.** Case schools and data forms.

	School A	School B	School C	School D
Funder	Public	Private Non-profit	Private For-profit	Private For-profit
Grade	10–12	10–12	10–12	7–9
Data	One principal interview; four teacher interviews; school documents	One principal interview; four teacher interviews; school documents	One principal interview	Four teacher interviews; school documents



**Table 2.** Background of teacher participants.

Case School	Participant	Background
School A	Betty	Music teacher, working since 1998
	Lara	Biology teacher, working since 2002
	Lisa	English teacher, working since 1999
	Noah	History teacher, working since 2013
School B	Jerry	Math teacher, working since 1984
	Phoebe	Chemistry teacher, working since 1986
	Hannah	Chinese teacher, working since 1989
	Hila	English teacher, working since 1995
School D	Camila	Physics teacher, working since 1999
	Selina	Dancing teacher, working since 2018
	Zena	Chinese teacher, working since 2014
	Charles	Math teacher, working since 2016

Second, the analysis of power relation explores the interactions between different accountabilities by looking into the relations between their intended objectives and the strength of their influence on final decisions. It intends to detect the potential conflicts between different accountabilities and how they negotiate about the disputes.

Last, the analysis of practical impacts evaluates mainly the motivating effect of each type of accountability, identifying teachers' motivational status and change with regards to influences of different accountabilities according to SDT (Ryan & Deci, 2000). SDT distinguishes different types of motivation based on different reasons that drives action and argues that the motivational types matter in eliciting and sustaining people's efforts (Deci & Ryan, 1998). It distinguishes motivational status through a spectrum of internalization where there are several transitional stages from least intrinsically driven to mostly intrinsically driven, including motivation, external regulation, introjection, identification, integration, and intrinsic motivation (Ryan & Deci, 2000). In addition to direct self-report motivational status, teachers' positive or negative comments also give clues about their attitude toward different accountabilities and the extent of their internalization.

In practice, teacher appraisal measures can be highly sensitive to and dependent on context, and the nominally identical practices may turn out to be completely different in accountability goals. For instance, "class observation," a practice commonly seen in both public and private schools, can be either administrative, professional, or market-oriented depending on "who observes," "for what purpose," and "in what way." When dominated by administrative accountability, class observation is usually conducted by school administrators with quantitative rubrics to score teachers'

performance as evidence to inform personnel management decisions, leaving little space for professional dialogue with teachers about development. On the contrary, class observation led by professional accountability, although may engage administrators, is mainly conducted by colleagues to encourage exchange and mutual learning in forms of research project, lesson study, or professional dialogue, during which teachers make their own roles and seek recognition from the professional learning community. In other circumstances like “school open day” in private schools, “class observation” can also be oriented by market accountability since the original intention of opening class to parents is to attract students in the competitive market by making a good impression on parents and allowing them to judge teachers for their own good. Thus, data analysis needs to go beyond the surface practices to detect the meaning of specific merit pay policy by teasing out the primary beneficiary, power relations, and practical impact behind similar merit pay regulations.

### **Public school: Autonomous teachers under professional-led accountability system**

School A as a public school, its merit pay scheme was largely subject to government policies with limited autonomy for school-based adaptations. About 70% of public-school teachers’ remuneration was made up of scheduled “national pay” and “subsidy,” which was solely dependent on teachers’ “professional title” and “working experience” (Bureau of Education Shenzhen, 2015; Pang et al., 2019). The total of post-salary and level salary was also named as “basic merit pay” since the 2009 merit pay reform, which was covered by fiscal appropriation allocated by district and municipal government.

Among the determinants of the “basic merit pay,” the professional title appraisal was the most decisive one. It ranked teachers into five leveled titles starting from “clerk,” “assistant,” “associate,” “advanced” to “senior.” Since the integration reform of appraisal and employment in 2015, the professional title appraisal lower than “advanced” level (included) has been authorized for public schools (Bureau of Human Resource and Social Security Shenzhen, 2015). Triennially, public schools shall convene a “Professional Title Appraisal Committee” consisting of holders of higher titles as the authoritative gatekeeper to evaluate candidates and determine the results via ballot. The appraisal of titles under “advanced” level (excluded) was mainly based on application materials, while candidates for “advanced” level need an extra oral defense.

The quantitative criteria were set up by the provincial government in *The Evaluation Standards for School Teachers in Guangdong Province (Pilot)* (Department of Human Resource and Social Security Guangdong, 2016), defining teachers’ performance by results of annual evaluation, academic publication, and cumulated awards of various teaching contests, in addition to the threshold requirements of degree and working experience. Students’ test scores were completely excluded

from the rubrics of the professional title appraisal, indicating a broader recognition of pedagogic values and its acknowledgment of teachers' stance.

In forms of a career ladder, the professional title appraisal was basically in accordance with the professional accountability for its committee-based organization and the profession-oriented design of rubrics. When asked about comments on the professional title appraisal, teachers were basically content with it for the abundant opportunities for professional growth and the effectiveness of sorting out excellent teachers.

*We used to have teaching contests triennially, but now it's annually. We can either choose to take part in a lesson study contest by our own, or to train students for performative contests. Opportunities are open and fair with encouragement for teachers under 35 years old to participate ... To excel in the professional title appraisal, you have to be really outstanding. All excellent teachers have finally passed.*

Betty, 2–6

*There is nothing upsets me in the current system of professional title appraisal. I think it's just the feature of being a teacher that much of our work is immeasurable and there is hardly any system that can accurately define good teachers from the poor. The current system allows teachers to be excellent in various ways with considerable autonomous space.*

Lara, 20

As shown in the excerpts, Betty appreciated the professional title appraisal for facilitating teachers' professional growth, and Lara recognized it for the autonomous space retained for teachers. It could be inferred from the positive comments that teachers in School A had internalized the professional codes to a great extent and mainly relied on intrinsic motivations to take the opportunities for professional growth or stick to their own beliefs as empowered by the autonomous space.

The rest 30% of public-school teachers' remuneration was named as "incentive merit pay," and its distribution was determined by school-based performance appraisal. As shown in Table 2, the performance appraisal measured teachers' teaching workload, administrative post, attendance, and related awards according to School A's "Distribution Plan of the Incentive Merit Pay." The measurement converted teachers' work into abstract commensurable units counted as multiple Xs, which usually equaled to 100 yuan and adjusted with the total of "incentive merit pay" (Table 3).

In the conversion of teachers' workload, the performance appraisal encouraged test preparation by attaching extra coefficient to academic subjects, especially the graduating grade, and including morning reading and tutoring class as standard lesson hour. While it sounded reasonable that teachers of academic subjects have extra workload from correcting students' assignment and test-oriented training, teachers of non-academic subjects still felt mistreated and tended to withdraw efforts at work secretly to reach psychological balance.

**Table 3.** The distribution plan of incentive merit pay in School A.

Requirement	Measurement
Dutifulness Bonus (12X)	Deduction: 1-4X Deduction: 1-4X Deduction: 2-6X
Full attendance bonus (8X)	None: two or more leaves 4X; one leave in a month 8X; full attendance
Extra lesson bonus (1-12X)	LH = (LP + LC × C) × N Award: (LH - 13) × X
Post Allowance (2-18X)	Subjects without assignments: N = 1 Subjects with assignments: N = 1.1 Academic subjects on graduating grade: N = 1.15 Chinese and English morning reading = 1 LH Tutoring class of academic subjects = 1 LH Group Head of Lesson Plan: 2X (less than five members) or 3X (six or more members) Group Head of Pedagogic Research: 6X (less than 16 members) or 8X (16 or more members)
Achievement Bonus	Head Teacher: 7X + 500 (national allowance) CCYL Secretary or Vice Dean of Department: 14X Grade head: 14X (average grade) or 16X (graduating grade) Dean of Department: 16X Vice principal: 18X The second post allowance counted in half School-level award: 1-5X District-level award: 3-10X Municipal-level award: 5-20X Provincial-level award: 10-30X National-level award: 20X +

Note. LH = lesson hour; LP = lesson plan; LC = lessons per class per week; CCYL = Chinese Communist Youth League; C = class; N = coefficient; X = unit of bonus, usually equals 100 yuan.

*Indeed, there is gap in income for the same teaching workload. For example, Physics teachers will receive more (pay) than Music teachers for one lesson ... We used to be quite negative about such unjust stipulations. Yet, would you quit or get along with it? If you wish to stay, then you have to get along with it. Under such conditions, we, I mean many Music teachers, chose to stop bothering ourselves with it. Some teachers like me just tend to do as much work as paid or even become lazing away.*

*Betty, 14–15*

As indicated by Betty, School A's performance appraisal failed to reconcile the dispute among teachers and inevitably crippled teachers' willingness to commit themselves. For teachers who found the stipulation unfair, they tended to view it as a compromise for retaining in School A and would withdraw part of their efforts as tacit resistance until they reached a psychological balance.

The conflict embedded in the performance appraisal indicated that the dominating stance was to maximize the gain in test scores as demanded by parents who represented the market accountability and to facilitate surveillance of teachers' input efforts for the good of school administrators. While teachers' full attendance brought them a monthly bonus of 800 yuan, two times of personal leaves or absence from school activities without proper excuse would result in punitive deduction in full. The surveillance of administrative accountability exerted "external regulation" upon teachers and drove their motivation toward the extrinsic direction.

*The introduction of performance appraisal and merit pay has led to fierce debate with endless meetings. Young teachers want to claim more because they are doing more work, while teachers of higher titles argue that the total of merit pay appropriated to school is dependent on numbers of holders of higher professional titles. As a result of futile negotiation, no one is happy with the plan.*

*Lisa, 28–29*

As shared by Lisa, teachers were distracted by the merit pay and the performance appraisal from teaching to the debate about the measurement. Whatever the result, teachers had been induced to pay more attention to extrinsic stimulus rather than maintaining the intrinsic rewards in teaching itself, causing an interruption to the autonomous working environment laid by the professional title appraisal.

In sum, School A's accountability system was featured by professional accountability as the primary force to normalize teachers with professional codes, while the administrative accountability functioned as the secondary force to monitor teachers' observable input efforts. Teachers' motivational status in relation to professional accountability was largely intrinsically driven in an autonomous atmosphere, while the administrative accountability had caused particular externalization of teachers' motivational status due to the controversies on merit pay. Market accountability was least influential since School A did not rely on tuition income.

## **Non-profit private school: Compliant teachers under administrative-balanced accountability system**

As a non-profit private school, School B was founded by an overseas Chinese entrepreneur as a philanthropy. The tuition fee was kept as low as 6,500 yuan per semester when public school students enjoyed 20,000 yuan per capita funding and for-profit private schools charged more than 20,000 yuan. Therefore, teachers' remuneration in School B was also limited, ranging from 5,000 to 10,000 yuan monthly in 2019, when public school teachers' remuneration started from 11,420 yuan and the average in for-profit private schools was above 10,000 yuan.

Principal B had been working in the Institute of Educational Sciences in Shanghai for years before he came to School B in Shenzhen, hence borrowed many common strategies from public schools. He was also personally against merit pay, emphasizing that teaching was incomparable to industrial manufacture, and a considerable portion of teachers' efforts was unquantifiable (Principal B, 3). Under Principal B's leadership, School B's merit pay scheme was mainly featured by administrative accountability balancing the tension between market accountability and professional accountability.

As borrowed from public school, the measurement of teachers' workload in School B was also structured with a uniform base pay guaranteed by a minimum requirement of lessons per week and extra pay for extra lesson hours exceeding the minimum requirement. However, teachers were not motivated as expected because of the meager income. To help promote teachers' income within the given structure and resources, Principal B lowered the requirement of minimum lesson hours from 10 to 5 per week. Thus, a larger proportion of work was counted as "extra," and teachers' income got promoted without changing the original salary structure or aggravating teachers' burden. While the direct beneficiaries seemed to be teachers, the nature of extra pay for extra workload was administrative accountability in essence since such adaption involved neither professional requirements nor inclusion of teachers in the decision-making process, and the primary concern was, in fact, to retain teachers from job-hopping with little implications for professional growth.

Excluded from the professional title system as a private school, School B lacked systemic professional accountability to encourage in-service professional development. To make up for the loss, Principal B borrowed a few strategies from public schools including subsidies scheduled by teachers' working experience and professional title, school-based teaching contests, lesson studies, class observations, personal yearly development planning, training of research methods, etc. With hopes to better motivate teachers, pecuniary incentives were attached to the measurement of teachers' performance related to professional activities.

In terms of gains in test scores, Principal B's effort had paid off with the passing rate of College Entrance Examinations raised from 27.4% to 99.51% in the past eight years. Yet, in terms of teachers' motivational status, Principal B might only be getting half the result with twice the effort.

*In terms of teaching contests, I would just do it if required by administrators. If were not the regulations, no one would take initiative. The pecuniary incentives hardly matter. It's just meaningless with such a low ceiling of bonus. As an old teacher of thirty years' experience, I could only rely on my own enthusiasm for education.*

Jerry, 4

As shared by Jerry, teachers in School B tended to be indifferent to the meager bonus and perceived the professional activities as external administrative tasks, indicating a limited extent of internalization and basic level of motivational status. Although Jerry was still intrinsically motivated by his enthusiasm for education, this drive was independent from the merit pay scheme, while the merit pay scheme was actually diminishing it rather than maintaining it. Such findings echoed with the crowding-out effect that people accustomed to intrinsic motives may experience decline in motivation when the incentives were changed from intrinsic to extrinsic. In other words, the effect of merit pay actually counteracted the contributions of Principal B's professional leadership to a certain extent given the limited financial resources, resulting in the alienation of professional activities toward administrative tasks as perceived by teachers.

In the meanwhile, despite Principal B's personal disapproval of merit pay, School B still relied on pecuniary incentives to maintain administrative control and to meet market demands.

As shown in Table 4, School B converted teachers' performance in "teaching routines," "professional activities," and "teaching quality" into several ranks attached with corresponding bonuses. Among these categories, "teaching routines" represented the stance of administrative accountability by regularly checking teachers' lesson plans and feedbacks on students' assignments, monitoring teachers' class management by administrators patrolling around the teaching buildings and recording the class discipline, and demanding for full attendance and absolute compliance with temporary work arrangements such as invigilation and proposition of test papers. "Professional activities" were originally designed to complement the professional accountability but ended up as auxiliary to administrative accountability due to the crowding-out effect in a limited financial condition, shifting the motivational force from intrinsic pursuit of pedagogic values to compliance with administrative regulations. "Teaching quality" was measured based on students' test scores in every mid-term and final examination, responding to parents' fervent demand of progress in academic performance. As a private school with heavy dependence on tuition income, School B had to respond to market accountability as much as it could.

**Table 4.** The merit pay scheme in School B.

Category		Rank			
		A	B	C	D
Teaching routines	Full attendance	800			
	Teaching material	350	300	250	–
	Class management	350	300	200	–
	Lesson plan	250	150	100	–
	Assignment correction	250	200	150	–
	Invigilation and proposition	300	100	50	–
	Morality supervision	400	300	200	–
Professional activities	Lesson study	300	200	150	–
	Publication	400	300	200	–
	Class observation	250	200	100	–
Teaching quality		1,200	1,000	800	500

In addition to the bonus included in the ranking system, School B actually retained another lump-sum bonus to reward teachers whose students outperformed the others in the College Entrance Examinations. With an average of 4,000 yuan, subject teachers whose students scored higher than the municipal average got an extra coefficient of 1.2 or 1.3, while those whose students scored lower than municipal average suffered from a deduction of about 0.2 or 0.3 coefficient (Principal B, 6–7). Numbers of students getting into first-tier universities brought teachers extra coefficient mediated by their backgrounds. One admission to first-tier universities from key classes brought an extra coefficient of 1, while one admission to first-tier universities from average classes brought an extra coefficient of 1.6 (Principal B, 6–7). Thus, the highest amount of the lump-sum bonus for CEE could reach as high as 20,000 yuan according to past cases.

In sum, School B's merit pay scheme was centered on administrative accountability, which proactively integrated professional norms to maintain teachers' professional ethics, and compromised to market accountability to hold teachers accountable to students' test scores. With the different accountabilities balanced within the administrative structure, teachers followed Principal B's leadership in teaching and research with basic-level intrinsic motivations. Meanwhile, the crowding-out effect of pecuniary incentives also indicated that it was the "nature of teaching" that rejected the ideology of merit pay.

### **For-profit private school: Struggling teachers under market-dominated accountability system**

*The Amendments of the Promotion Act of Minban Education* (2016 Amendments) officially put forward the "Classified Management" of private schools, in which for-profit private schools



were allowed to extract profits from the surplus for their investors but would be receiving less government supports. In principle, schools at compulsory education stage cannot be registered as for-profit. But in reality, many education groups were operating K-12 schools like School C and School D, attracting students from compulsory education stage and claiming profits from senior secondary faculty. School C charged 31,000 yuan for each semester for upper secondary in 2020 and School D charged 76,000 yuan for each semester for lower secondary in 2020, much higher than the 6,500 yuan charged by School B.

As part of a larger education group, School C has abundant financial resources. Teachers' monthly remuneration in School C ranged from 10,000 to 15,000 yuan, which was much higher than that in non-profit private schools like School B. Principal C appreciated the rationale and motivating function of merit pay, so he designed a typical merit pay system named "post accountability and goal management system" (PAGMS). It was featured with the hierarchical objective structure and the rated bonus based on the extent of fulfillment. At the beginning of each academic year, teachers were supposed to set up objectives in every aspect of evaluated performance, which were subordinate to the objectives of their superior administrators. About 60% fulfillment of objectives equaled 80% bonus, and the upper limit was 120% bonus, which required 100% fulfillment. The average bonus could reach as high as 120,000 yuan for an academic year.

The indicators of PAGMS were mainly outcome-based and at the service of market accountability. For instance, objectives of teaching outcome were defined by four indicators: (1) students' mean scores in standardized tests, (2) the ratio of high achievers versus underachievers, (3) students' awards in various competitions, and (4) the final outcome of the College Entrance Examinations. Moreover, the PAGMS imposed the indicator of "student stability" on teachers' work, holding teachers accountable for the rate of students retaining at School C. It explicitly revealed the dominance of market accountability, which treated teachers more like marketing personnel rather than professionals.

Unlike School C's goal management system, School D responded to market accountability by including parents in school management and teacher appraisal. Parents of School D were constantly leading the flag-raising ceremony, parents' academy, festival celebrations, mixed-age family routines, and the basketball club management. Moreover, they can observe teachers' teaching directly in classrooms, check teachers' work records on a digital platform, and debrief teachers about their teaching performance in one-to-one consultative meetings.

While teachers in School D were closely monitored by parents, their work was labor intensive. School D de-constructed teachers' work into a few independent posts (like family tutor, mental health teacher, head teacher, group head, subject head, and faculty head), with packaged salary ranging from 2,000 to 6,000 yuan for each, making it insufficient to make a living by only taking one post but considerably attractive to take two or more posts concurrently. Although it

was also common in public schools like School A that teachers may take additional administrative posts, the income gap brought by administrative workload ranged only from 200 to 1,800 yuan in School A with a minimum income as high as 11,420 yuan, while the gap ranged from 2,000 to 6,000 yuan in School D with a minimum income as low as 6,000 yuan. As a result, teachers in School D had to take two or more posts concurrently to reach a satisfactory income of around 10,000 yuan.

Besides the monthly salary piled up by intensive work, School D also adopted an annual bonus with delayed payment to retain teachers. Teachers had to remain at School D until next semester to get the annual bonus for the last academic year, otherwise, they would lose this bonus if they quit the job voluntarily. The amount of annual bonus was ranked into three tiers: 10,000, 15,000, and 20,000 yuan. Yet, its determination was controlled by administrators in a black box, which excluded teachers and lacked explicit standards for the ranking. Teachers had to vote by feet if they were not going to deal with it, and so did they. Many teachers had already left for another school or another occupation as reported by interviewees, and School D had to keep recruiting more new teachers to maintain its faculty. Such instability of teaching team might be attributed to the anti-democratic administrative accountability, as well as the irreconcilable conflicts between professional ethics and market accountability.

*Once we failed a student in an English Sing Competition. Soon after that we received a text message from the students' parent, saying that 'How come my kid get failed? He used to win various prizes in other off-campus competitions. How come he failed in yours?' Then another parent also came to criticize the organization of the competition. Under the oppressive pressure, we had no choice but to let the student try again and again. We changed the rules for a parent who was reluctant to accept the fact that her son was not winning prizes ... Teachers originally wish to teach them well and make a difference. But you need to scruple about parents. When a student makes a mistake but his parent wants no more criticism for him, teachers cannot help but step backward. We don't know what to do under parents' interventions, and students become unscrupulous since they know that you can't actually restrain them.*

*Selina: 17-2-13*

As reported by Selina, parents' demand for individualized service was in essence self-interested preferential treatment rather than pedagogic concern to teach students according to their aptitude. Vulnerable to parents' meddling demands, teachers were inevitably frustrated by the irreconcilable conflict between market and professional accountability.

To sum up, the accountability structure of for-profit private schools was featured with collusion between market accountability and administrative accountability. Specified by the case of School D, market accountability might constantly interrupt teachers' work since parents' demands were mainly self-interested. Lacking supportive force, teachers in for-profit schools were vulnerable to

external requirements from market and administrative accountabilities and were struggling among the irreconcilable conflicts between market accountability and their own pedagogic concerns. Under such accountability system, teachers' motivational status was least internalized with inclination to leave for a better school or occupation. The variations of merit pay forms between School C and School D also indicated that it was the market-oriented accountability structure that matters rather than the technical structure of merit pay.

## **Conclusion and discussion**

In this study, we fully recognized the variations of merit pay schemes and attempted to capture the nuances among public, non-profit private, and for-profit private schools by analyzing the accountability structure. This analytical lens made comparison between different school contexts possible and meaningful and helped to expose the power relation behind the trivial measurements. We also chose teachers' motivational status instead of students' academic progress to depict the impacts of merit pay schemes on the process and conditions of teachers' work, not merely the results. Thus, teachers' experience of potential conflicts became visible and their voice received attention that they deserved. Yet, the limitation of this study was also apparent that it failed to include more direct viewpoints and experiences from parents and school administrators or observation of interactional situations to enable a thicker description. Despite that, the findings still generated profound implications about the ongoing debate between the "nature of teaching" hypothesis and the "political obstruction" hypothesis.

To the contrary of the "political obstruction" hypothesis, our findings suggested that the more dependent on the market, the more intense the conflict between external accountabilities and teachers' internalized professional concern. The struggles of teachers in School D indicated that they were aware of the interference brought by the market accountability but were still half-way aware of the essence of market-oriented "teaching subject" (Ball & Olmedo, 2013). A similar conflict was also embedded in the discriminative measurement of workload in School A, but the discouraging effect was largely counteracted thanks to the dominant professional accountability. The crowding-out effect of pecuniary incentives in School B also supported the "nature of teaching" hypothesis that merit pay would probably damage teachers' intrinsic motivation due to their inherent incompatibility, despite its contribution to extrinsic motivation. Thus, our findings suggested timely pause of merit pay reform in public schools and hopefully a new reform based on the nature of teaching.

## **Contributorship**

Hejia Shi contributed to the data collection and the first draft of this article. Dan Wang contributed to the revision of theoretic framework and was closely engaged in data analysis and presentation.

## Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Ethical statement

A written consent form was signed by interview participants before data collection and a copy of interview transcript was proofread by interview participants afterward. The study was approved by the Human Research Ethics Committee of the University of Hong Kong (reference number: EA1811027).

## Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This article is based on the corresponding author's doctoral study financed by Post-Graduate Scholarship of the University of Hong Kong (HKU).

## ORCID iD

Hejia Shi  <https://orcid.org/0000-0002-6250-1658>

## References

- Al-Ubaydli, O., & Lee, M. S. (2012). Do you reward and punish in the way you think others expect you to? *The Journal of Socio-Economic*, 41(3), 336–343. <https://doi.org/10.1016/j.socec.2012.02.002>
- Ball, S. J., & Olmedo, A. (2013). Care of the self, resistance and subjectivity under neoliberal governmentalities. *Critical Studies in Education*, 54(1), 85–96. <https://doi.org/10.1080/17508487.2013.740678>
- Ballou, D. (2001). Pay for performance in public and private schools. *Economics of Education Review*, 20(2001), 51–61. [https://doi.org/10.1016/S0272-7757\(99\)00060-6](https://doi.org/10.1016/S0272-7757(99)00060-6)
- Ballou, D., & Podgursky, M. (1993). Teachers' attitudes toward merit pay: Examining conventional wisdom. *Industrial & Labor Relations Review*, 47(1), 50–61. <https://doi.org/10.1177/001979399304700104>
- Benabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *Review of Economic Studies*, 70(3), 489–520. <https://doi.org/10.1111/1467-937X.00253>
- Bowles, S., & Polania-Reyes, S. (2012). Economic incentives and social preferences: Substitutes or complements? *Journal of Economic Literature*, 50(2), 368–425. <https://doi.org/10.1257/jel.50.2.368>
- Bureau of Education, Shenzhen. (2015). *The pilot for deepening the reform of the schoolteachers' professional title institution* [in Chinese].
- Bureau of Human Resource and Social Security, Shenzhen. (2015). *The pilot for deepening the reform of the professional title system of schoolteacher* [in Chinese].
- Cacioppe, R., & Mock, P. (1984). A comparison of the quality of work experience in government and private organizations. *Human Relations*, 37(11), 923–940. <https://doi.org/10.1177/001872678403701104>
- Cooper, S., & Cohn, E. (1997). Estimation of a frontier production function for the South Carolina educational process. *Economics of Education Review*, 16, 313–327.

- Crewson, P. E. (1997). Public-service motivation: Building empirical evidence of incidence and effect. *Journal of Public Administration Research and Theory*, 7(4), 499–518. <https://doi.org/10.1093/oxfordjournals.jpart.a024363>
- Darling-Hammond, L. (2004). Standards, accountability, and school reform. *Teachers College Record*, 106(6), 1047–1085. <https://doi.org/10.1111/j.1467-9620.2004.00372.x>
- Dee, T., & Keys, B. J. (2004). Does merit pay reward good teachers? Evidence from a randomized experiment. *Journal of Policy Analysis and Management*, 23, 471–488.
- Department of Human Resource and Social Security, Guangdong. (2016). *The evaluation standards for schoolteachers in Guangdong province (Pilot)* [in Chinese].
- Dessler, G. (2011). *Human resource management* (12th ed.). Pearson Prentice Hall.
- Dielmann, K. M. (2012). Merit pay. In W. J. Rothwell & R. K. Prescott (Eds.), *The encyclopedia of human resource management* (pp. 314–316). Pfeiffer.
- Falk, A., & Kosfeld, M. (2006). The hidden costs of control. *American Economic Review*, 96(5), 1611–1630. <https://doi.org/10.1257/aer.96.5.1611>
- Fan, X., & Fu, W. (2011). Merit pay reform in compulsory education schools: Background, effectiveness, problems and solutions—Investigation of 4 provinces and 32 counties in mid-China [in Chinese]. *Journal of Huazhong Normal University (Humanities and Social Sciences)*, 50(6), 128–137.
- Frey, B. S. (1997). *Not just for the money*. Edward Elgar Publishing.
- Frey, B. S., & Osterloh, M. (2002). *Successful management by motivation. Balancing intrinsic and extrinsic incentives*. Springer.
- Frey, B. S., & Osterloh, M. (2013). Motivation governance. In A. Grandori (Ed.), *Handbook of economic organization: Integrating economic and organization theory* (pp. 26–40). Edward Elgar.
- Fu, W. (2013). Efforts to establish financial security system to guarantee teachers' merit pay in compulsory education schools [in Chinese]. *Education & Economy*, (3), 9–14.
- Fu, W., & Zeng, X. (2010). Problems and solutions to teachers' merit pay reform in compulsory education schools—Based on investigation of counties in four mid-China provinces [in Chinese]. *Collection of Thesis of the Annual Conference of Educational Economics in China*.
- Gailmard, S., & Patty, J. W. (2007). Slackers and zealots: Civil service, policy discretion, and bureaucratic expertise. *American Journal of Political Science*, 51(4), 873–889. <https://doi.org/10.1111/j.1540-5907.2007.00286.x>
- Gao, P., Peng, J., & Tan, L. (2007). Research on the adjustment of county-level teachers' union in protecting teachers' rights in rural schools [in Chinese]. *Journal of Anhui Agricultural Science*, 35(3), 855–856.
- Gius, M. (2013). The effects of merit pay on teacher job satisfaction. *Applied Economics*, 45(31), 4443–4451. <https://doi.org/10.1080/00036846.2013.788783>
- 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.
- Hannaway, J. (1992). Higher order skills, job design, and incentives: An analysis and proposal. *American Educational Research Journal*, 29(1), 3–21. <https://doi.org/10.3102/00028312029001003>
- He, X. (2019). Questions about school teachers' merit pay system in China [in Chinese]. *China Economics of Education Review*, 4(5), 3–16.

- Hood, C. (1991). A public management for all seasons? *Public Administration*, 69(1), 3–19. <https://doi.org/10.1111/j.1467-9299.1991.tb00779.x>
- Huang, B., Cheng, H., He, P., & Du, P. (2021). Pecuniary income, working conditions and rural teachers' intention of rotation and resignation—Empirical research based on data from Chongqing and Ningxia [in Chinese]. *Education & Economy*, 37(5), 77–88.
- Lah, T. J., & Perry, J. L. (2008). The diffusion of the civil service reform act of 1978 in OECD countries: A tale of two paths to reform. *Review of Public Personnel Administration*, 28(3), 282–299. <https://doi.org/10.1177/0734371X08319950>
- Lai, D. (2013). Growing salary of teachers, growing regional disparity—Regional difference analysis of teachers' salary in China [in Chinese]. *School Administration*, (2), 36–38.
- Li, K. (2013). Analysis on policy implementation of performance wage in different types of schools in compulsory education [in Chinese]. *Educational Research*, (5), 46–52.
- Li, X. (2010). Status quo, issues and suggestions of teacher performance pay at compulsory schools [in Chinese]. *Journal of Shandong Agricultural Administrators' College*, 27(3), 170–171.
- Lidenberg, S. (2001). Intrinsic motivation in a new light. *Kyklos*, 54(2–3), 317–342. <https://doi.org/10.1111/1467-6435.00156>
- Liu, H. (2012). The issue of teacher performance pay in solving unequal payment within the same city [in Chinese]. *Construction of Faculty Forum*, (6), 84.
- Liu, S., Xu, X., & Stronge, J. (2016). Chinese middle school teachers' preferences regarding performance evaluation measures. *Educational Assessment, Evaluation and Accountability*, 28, 161–177. <https://doi.org/10.1007/s11092-016-9237-x>
- Meng, W., & Yuan, Q. (2014). Characteristics of primary and middle school teachers' performance evaluation system—Literature analysis of multiple cases study based on teacher performance evaluation plan in X city [in Chinese]. *Educational Measurement and Evaluation*, (8), 26–37.
- Ministry of Education of China. (2008). *The guiding opinions on implementing merit pay in compulsory education schools* [in Chinese].
- Mumane, R., & Cohen, D. (1986). Merit pay and the evaluation problem: Why most merit pay plans fail and few survive. *Harvard Educational Review*, 56, 1–17. <https://doi.org/10.17763/haer.56.1.l8q2334243271116>
- Noe, R. M., Hollenbeck, J. R., Gerhart, B., & Wright, P. M. (2011). *Fundamentals of human resource management* (4th ed.). McGraw-Hill/Irwin.
- Paine, L., & Fang, Y. (2006). Reform as hybrid model of teaching and teacher development in China. *International Journal of Educational Research*, 45(4), 279–289. <https://doi.org/10.1016/j.ijer.2007.02.006>
- Paine, L., & Ma, L. (1993). Teachers working together: A dialogue on organizational and cultural perspectives of Chinese teachers. *International Journal of Educational Research*, 19(8), 667–778. [https://doi.org/10.1016/0883-0355\(93\)90009-9](https://doi.org/10.1016/0883-0355(93)90009-9)
- Pang, L., Yang, X., & Jin, Z. (2019). Difficulties, impacts and policy implications of rural teachers' professional title appraisal [in Chinese]. *Teacher Education Research*, 31(1), 31–36.
- Pham, L. D., Nguyen, T. D., & Springer, M. G. (2021). Teacher merit pay: A meta-analysis. *American Educational Research Journal*, 58(3), 527–566. <https://doi.org/10.3102/0002831220905580>

- Ruffini, R., Modarelli, G., Sferrazzo, R., & Turri, M. (2020). Is merit pay changing ethos in public administration? *Cogent Business & Management*, 7(1), 1724703. <https://doi.org/10.1080/23311975.2020.1724703>
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54–67. <https://doi.org/10.1006/ceps.1999.1020>
- Tang, Y., & Hu, Y. (2013). Institutional structure design of teachers' salary in compulsory education schools in China—Perspectives of economics and management [in Chinese]. *Teacher Education Research*, (4), 20–25.
- Wang, B. (2005). Teacher appraisal methods and correspondent subjects [in Chinese]. *Teacher Education Research*, 17(1), 42–50.
- Wang, L., Lai, M., & Lo, L. N.-K. (2014). Teacher professionalism under the recent reform of performance pay in Mainland China. *PROSPECTS*, 44(3), 429–443. <https://doi.org/10.1007/s11125-014-9315-0>
- Wang, W. (2007). Evaluation of 2+2 alternative teacher performance appraisal program in Shanxi, People's Republic of China. *Teaching and Teacher Education*, 23(7), 1012–1023. <https://doi.org/10.1016/j.tate.2006.05.005>
- Warren, A. N., & Ward, N. A. (2019). “It didn't make me a better teacher”: Inservice teacher constructions of dilemmas in high-stakes teacher evaluation. *School Effectiveness and School Improvement*, 30(4), 531–548. <https://doi.org/10.1080/09243453.2019.1619185>
- Weibel, A., Rost, K., & Osterloh, M. (2007). *Crowding-out of intrinsic motivation-opening the black box*. <https://dx.doi.org/10.2139/ssr.957770>
- Weibel, A., Rost, K., & Osterloh, M. (2010). Pay for performance in the public sector—benefits and (hidden) costs. *Journal of Public Administration Research and Theory*, 20(2), 387–412. <https://doi.org/10.1093/jopart/mup009>
- Williams, J., & Engel, L. (2012). How do other countries evaluate teachers? *Phi Delta Kappan Magazine*, 94(4), 53–57. <https://doi.org/10.1177/003172171209400414>
- Woessmann, L. (2011). *Merit pay international: Countries with performance pay for teachers score higher on PISA tests. (Program for International Student Assessment)*. *Education Next*, 11.
- Xiong, B. (2009). *Teacher performance pay limited by educational system* [in Chinese]. <http://edu.people.com.cn/GB/10529384.html>
- Xu, X., & Yang, T. (2008). Comparison of teachers' union and teachers' congress in schools [in Chinese]. *Shandong Trade Unions' Tribune*, 14(6), 42–43.
- Yuan, K., Le, V., McCaffrey, D. F., Marsh, J. A., Hamilton, L. S., Stecher, B. M., & Springer, M. G. (2013). Incentive pay programs do not affect teacher motivation or reported practices: Results from three randomized studies. *Educational Evaluation and Policy Analysis*, 35(1), 3–22. <https://doi.org/10.3102/0162373712462625>
- Zhang, X., & Ng, H. (2017). An effective model of teacher appraisal: Evidence from secondary schools in Shanghai, China. *Educational Management Administration & Leadership*, 45(2), 196–218. <https://doi.org/10.1177/1741143215597234>
- Zhang, Z. (2010). Where the “teacher performance evaluation plan” can go—A survey study of the starting point, process and results of teacher evaluation plan at Z school [in Chinese]. *Journal of the Chinese Society of Education*, (5), 19–22.

- Zhao, H., Hui, X., & Fu, C. (2011). Investigation of the implementation of merit pay schemes in compulsory education schools in China [in Chinese]. *Theory and Practice of Education*, 31(10), 24–27.
- Zhao, J., & Liu, M. (2017). Macro mechanism of teacher salary system and implementation effects at localities [in Chinese]. *Peking University Education Review*, 15(2), 2–16.