# Arts Education Policy Review

# Merit Pay and the Music Teacher

Kenneth Elpus [a]

[a] University of Maryland , College Park, Maryland, USA
Published online: 23 Aug 2011.

PLEASE SCROLL DOWN FOR ARTICLE

Routledge
Taylor & Francis Group

# Merit Pay and the Music Teacher

Kenneth Elpus

*University of Maryland, College Park, Maryland, USA*

Current proponents of education reform are at present seeking to fundamentally change the system of teacher compensation by eliminating the traditional single salary schedule and instituting a merit pay system that directly links teacher pay to student achievement. To date, the scholarly literature in music education has been silent on the subject of teacher compensation reform. This article reviews the political arguments and empirical evidence on teacher merit pay while considering these reforms' potential deleterious effects on music educators. After examining the potential pitfalls of a merit pay system for music educators, I propose one possible framework for evaluating music teachers in a merit pay system.

*Keywords:* education reform, music, music education, teacher compensation

Political calls for educational policy reform have a long history in the United States. For three decades, federal education policy in the United States has largely focused on proposals seeking to reform current educational practice (Vinovskis 2009). Initially an effort of the Republican Party and other like-minded conservative thinkers, modern educational reform since the mid-1980s has focused on breaking the longstanding traditions of the nation's education system that these groups believed had led to a relaxation of curricular rigor in the schools. Although this drive was first launched by proponents on the right side of the political spectrum, a growing cadre of centrist and left-leaning members of the Democratic Party have latched onto and support many of these reform efforts (Ravitch 2010). Currently led by the Democrats for Education Reform (DFER), these reformers are held in high esteem by the Obama administration, and many of their proposals have been codified in Race to the Top (RTTT), the $4.35 billion federal education reform inducement enacted as part of the 2009 American Recovery and Reinvestment Act (Goldstein 2009). Often rooted in the same economic thought that dominates the conservative movement, many of these reforms seek to apply analogues of free-market principles to education. Standards-based reform, school choice reform, and school accountability reform are each undergirded by the notion that schooling in America will be improved by removing education from the purview of local government and placing it in a self-regulating educational marketplace within a federally sanctioned framework that allows parents to vote with both their feet and their dollars.

Most education reform proposals have become contentious political issues, with proponents and critics vehemently debating their effectiveness, utility, and necessity in addressing a variety of deficiencies in the current system. These arguments have included debates on inequitable educational opportunity and the production of graduates who are unready to enter the workforce, to name only two.

Of these market-based educational reform initiatives, perhaps none has been as contentious as teacher compensation reform. Broadly, "teacher compensation reform" refers to any attempt to deviate from the traditional method of remunerating teachers in which teachers' compensation rises monotonically with each additional year of experience and with the attainment of or significant progress toward accredited graduate degrees. This traditional method of compensating teachers is referred to as the *single salary schedule*, named because of the grid that commonly appears in a teacher's employment contract listing academic credentials in columns (sometimes called the "lane") and years of experience in rows (the "step"), with the single cell that collocates a teacher's step and column containing that teacher's annual salary. Under this system, regardless of grade level or subject taught, all teachers in a district are paid the base salary listed in the agreed-upon single salary schedule, which may be augmented by the teacher taking on additional responsibilities, such as directing co-curricular ensembles or coaching extracurricular sports.

Correspondence should be sent to Kenneth Elpus, Assistant Professor in Music Education, University of Maryland School of Music, 2110 Clarice Smith Performing Arts Center, College Park, MD 20742-1620, USA. E-mail: elpus@umd.edu

Springer (2009) offers a useful taxonomy of teacher compensation reform schemes that reformers hope will augment or replace the single salary schedule:

1. *Market-based pay*, in which teachers hired in difficult-to-staff subjects or difficult-to-staff schools are paid at higher rates than other teachers.
2. *Knowledge- and skills-based pay*, in which teachers earn higher compensation based on their efforts to complete meaningful professional development linked to greater student achievement.
3. *Career ladders*, in which a predefined path of advancement grants teachers additional pay and greater responsibility.
4. *Recruitment and retention awards*, which are given as signing bonuses at the start of a teacher's career and annual bonuses for continued service to the same district.
5. *Pay-for-performance* or merit pay, in which a substantial portion of a teacher's annual compensation is formulaically tied to "predetermined tasks or outcomes, or both, related to teacher and student behaviors" (5).

To some extent, there is at least small-scale precedent for each of these teacher compensation reforms—even for music teachers—in the United States. In many geographic areas, teachers in difficult-to-staff subjects or difficult-to-staff schools are offered either recruitment awards or student loan forgiveness (which can be classified as market-based pay, since portions of the loans are cancelled for each additional year of service). Music teachers in low-income schools are currently eligible for this kind of market-based pay for up to $5,000 in Stafford loans and up to 100 percent of Perkins loans. Music teachers' Perkins loans can also be cancelled for teaching in a state that has officially classified music as a teacher shortage area.[1] Knowledge- and skills-based pay structures are evident in many states that pay teachers annual incentives for becoming and remaining National Board–certified teachers (NBCT) by completing the initial two-year certification process and fulfilling the continuing certification requirements of the National Board for Professional Teaching Standards (NBPTS). Two such NBCT certificates, stratified by the age of students taught, are available in music (National Board for Professional Teaching Standards [NBPTS] 2001; NBPTS 2010). Career ladders, featuring stratified certification levels with associated compensation incentives, have appeared in several states, notably in Tennessee during the 1980s (Dee and Keys 2004) and in Arizona at the present time (Buck and Greene 2011).

Of all the teacher compensation reforms available, perhaps none inspires more confidence and enthusiasm among free-market advocates than the pay-for-performance model. Indeed, merit pay schemes are now often perceived to be synonymous with teacher compensation reform, particularly given their central role in the Obama administration's RTTT

initiative. Music education scholarship has unfortunately been silent on the issue of teacher compensation—reformed or otherwise—though this silence is perhaps unsurprising, considering recent calls by Jones (2009) and Kos (2010) for more practical policy research and less theoretical policy scholarship in the field. In this article, I will present both sides of the scholarly debate over teacher merit pay, review empirical literature on the effectiveness of pay-for-performance schemes in education, discuss issues raised by the implementation of merit pay schemes in music education, and offer policy recommendations for the music education profession as it reacts to the nationwide push for a pay-for-performance model in teaching.

## POLICY ARGUMENTS IN SUPPORT OF MERIT PAY

As with any political or policy debate, numerous arguments have been put forth in support of or against pay-for-performance teacher compensation reform. In this section, I will present a brief overview of the arguments in favor of teacher merit pay that have appeared in the scholarly literature. In general, calls for teacher compensation reforms are "predicated on the argument that prevailing compensation practices provide weak incentives for teachers to act in the best interest of their students[,] and that inefficiencies arise from rigidities in current compensation policies" (Springer 2009, 1).

Podgursky (2009) most clearly lays out the free-market and economic principles that undergird support for both teacher compensation reform broadly and merit pay specifically. In describing the economic argument against the single salary schedule, Podgursky puts forth a "market clearing thesis" (75). Viewing the teacher workforce as a labor market, he explains:

> There is a saying in economics (the origin of which I do not know): "You can't repeal the law of supply and demand." By this, economists mean that if prices are not allowed to clear a market then some other mechanism will. . . . It is important to keep this idea of non-price clearing in mind in considering the effects of teacher salary schedules on the level and distribution of teacher quality(70).

To support his market-clearing thesis, Podgursky (2009) points out inefficiencies of the single salary schedule that could be improved through various teacher compensation reform schemes. He first compares easy-to-staff and hard-to-staff subject specializations: specifically, elementary teaching and high school chemistry teaching. Podgursky cites data showing that although both jobs pay the same salary under the single salary schedule, school districts often have difficulty filling secondary chemistry positions because of a lack of available applicants but generally have a glut of applicants

for elementary school positions. In this situation, principals can choose the highest quality elementary applicant from a large pool but will often have to settle for any available chemistry teacher, regardless of quality. Podgursky thus concludes that "the market 'clears' in terms of quality rather than price" (72).

In applying the market-clearing thesis to the pay-for-performance model, Podgursky notes that a vast amount of research on the economics of education has shown that there is a great deal of variation in teachers' ability to raise student achievement under any feasible definition of achievement. He refers to this ability as "teacher quality" or "teacher effectiveness." Under the single salary schedule, teachers of varying quality are assumed by school districts to be of equal quality and thus receive the same wage, a phenomenon that Weisberg et al. (2009) refer to as the "widget effect." Under the pay-for-performance model, Podgursky (2009) contends, teachers are motivated to immediately increase student achievement on the performance metric, and in the long term, the pay-for-performance model would recruit into the teaching workforce those teachers most able to thrive under the system and meet the measured goals. Without such incentives, the single salary schedule is inefficient, because there is no monetary reward for better teaching. To summarize, then, "equalizing teacher pay among teachers at different levels of effectiveness lowers the overall quality and performance of the teaching workforce. Once again, the market clears in terms of quality rather than price" (76).

Other than the free-market principles explained by Podgursky's market-clearing thesis and similar economic theory–based arguments, the primary philosophical claims in support of teacher merit pay plans center on first, the perceived disconnect between teachers' incentives and student performance, and second, the dissimilarity between teacher compensation and compensation policies in the private sector, where it is presumed that pay-for-performance is a routine compensation practice. Proponents of merit pay contend that under the single salary schedule, uniform pay schedules based on teacher-level inputs such as certification, graduate degrees, and years of experience do not incentivize teachers to focus on student outputs such as educational achievement and higher test scores (Lazear 2003). Ideologically, merit pay proponents believe that compensation for outputs is more efficient and desirable than compensation for inputs. Frequently cited as evidence in this argument is the accumulating, although not completely settled, body of research that suggests that student achievement is not correlated with either teachers' attainment of master's degrees or additional years of teaching experience beyond the initial gain attributable to experience in the first few years of teaching (Hanushek 2003; Rockoff et al. 2008). Springer notes that "the vast majority of variation in teacher effectiveness cannot be explained by observable teacher characteristics, that is, the type of teaching certificate held, level of education, licensing exam scores, and years of teaching experience" (2009, 10).

## POLICY ARGUMENTS OPPOSING MERIT PAY

Opponents of pay-for-performance teacher compensation schemes tend to reject the application of free-market principles to the educational context. Three main arguments have been used to oppose this model: first, the perceived emphasis on quantitative performance measures in private sector compensation does not accurately reflect reality; second, merit pay introduces "perverse incentives" that lead to teachers "gaming" the system, rather than marked student achievement gains; and third, certain merit pay schemes disrupt the inherently collaborative nature of teaching in deleterious ways.

In a review of private sector compensation policies, Adams and Heywood (2009) find that quantitative performance measures are rarer in the private sector than teacher compensation reformers contend. They also find that in the limited instances in which strictly quantitative pay-for-performance structures exist, they tend to make up a far smaller portion of the total compensation package than reformers assume. Adams and Heywood show that quantitative relationships between productivity and pay generally exist only in the finance, insurance, and real estate industries. While they find that many workplaces offer something referred to as a "bonus," these bonuses generally do not meet the definition of strict merit pay. Based on their analyses of several large economic datasets, Adams and Heywood ultimately conclude that "the suggestion that large shares of the private sector workforce have a tight formulaic relationship between earnings and performance is wrong" (57).

Rothstein (2009a; 2009b) invokes a social science maxim known as "Campbell's Law" to summarize the argument that quantitative performance measures lead to perverse incentives and unintended consequences. Campbell's Law, first proposed by social sciences research methodologist Donald T. Campbell, functions as a sort of Heisenberg uncertainty principle for the social sciences, particularly when social science methods are applied to the evaluation of public sector programs and services. Campbell proposes that "the more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor" (Campbell 1976, 54).

Using Campbell's Law as a theoretical grounding, Rothstein (2009a) shows that when strictly quantitative measures are used to evaluate performance in other fields, perverse incentives arise that lead to gaming of the system and unintended consequences, which he calls *goal distortion*. Rothstein argues that teacher compensation tied to test scores will incentivize the creation of narrower curricula—teaching to the test—that will raise test scores at the expense of a broader education that includes depth in nontested subjects such as music, thus distorting the true goals of education. He thus asks, "How much gain in reading and math scores is necessary to offset the goal distortion—less art, music,

physical education, science, history, character building—that inevitably results from rewarding teachers or schools for score gains only in math and reading?" (2009b, 98)

Other, more socially objectionable unintended consequences can also arise from a system in which teacher incentives are based on test scores. Using data from the Chicago Public Schools, Jacob and Levitt (2003) show that even a modest change in incentives can lead to instances of wholesale teacher cheating on standardized tests.

The final argument opposing teacher compensation merit pay schemes is that these models can create an unhealthy atmosphere of competition among teaching staff. This environment is particularly likely under "zero-sum" tournament-style merit pay plans, in which either the number of teachers who receive merit bonuses or the total expenditure on merit bonuses is fixed and limited from year to year. Podgursky and Springer summarize this criticism: "To a considerable extent, teachers work as members of a team. Introducing performance-related rewards at the individual teacher level might reduce incentives for teachers to cooperate and, as a consequence, reduce rather than increase school performance" (2007, 927).

Conceptually, critics contend that merit pay systems suffer from the fundamental flaw of failed social initiatives proposed by Scott in *Seeing Like a State* (1998). By necessity, merit pay schemes must make the work of teaching "legible" to outside authorities, who must be able to view classroom performance at a glance, distilled to a metric that is easily digestible from a high altitude. Scott argues that high modernist states must make social processes legible to a central authority to allow the state to effectively govern, and yet, in so doing, local knowledge and the virtues of heterogeneity are lost. In making teaching legible for merit pay purposes—much as standardized tests attempt to make learning legible—the art and skill of teaching a diverse group of students are reduced to an average test score metric that allows teachers to be "rated" by anyone, without ever needing to set foot in a classroom. Critics thus argue that using standardized measures of student learning to make teaching legible causes Campbell's Law to override the purported benefit of aligning teacher incentives and student performances, because the metrics themselves become corrupt indicators.

## EMPIRICAL EVIDENCE

Although much political enthusiasm surrounds teacher compensation reform in the United States, the single salary schedule is so entrenched in the American educational system that little completed domestic research has examined the effectiveness of pay-for-performance schemes in education. Springer (2009), however, reports that four U.S.-based randomized controlled trials of teacher performance pay schemes are currently underway, one of which examines individual teacher incentives and the rest of which exam-

ine group-based or hybrid group/individual performance pay incentives. One of those studies (Springer et al. 2010) has recently been completed. Although international research has been conducted on teacher incentives (Glewwe, Ilias, and Kremer 2003; Lavy 2002), in this section, I will present a brief overview of the empirical evidence that has examined teacher performance incentives in the United States.[2]

Figlio and Kenny (2007) conducted the first large-scale U.S.-based study attempting to discover a link between teacher performance incentives and student achievement. Recognizing that most nationally representative datasets in education cannot sufficiently link school compensation policies with student outcomes, they supplemented data from the National Education Longitudinal Study of 1988 (NELS) with their own survey of compensation policies used by the schools in the NELS sample, as well as data from the 1993 U.S. Department of Education Schools and Staffing Survey. Figlio and Kenny find that merit pay programs are linked to higher student test scores in schools where the merit pay program is individual (as opposed to group- or schoolwide) and small numbers of teachers receive the available bonuses. This association persists even when controlling for union status, state-level reform initiatives, and the school's status as public or private. Figlio and Kenny find this relationship between targeted merit pay programs and student achievement to be the strongest in schools serving low-income students. No relationship was found between group- or schoolwide incentives and student achievement. However, the researchers caution that they are unable to conclude unequivocally that the incentive program itself is causal: schools with teacher incentive programs during the period they studied may have had a higher overall school quality in difficult-to-measure ways that were unobserved in their dataset.

More recently, Goodman and Turner (2011) examined the impact of group-based incentive pay in the New York City public school system on teacher effort (as measured through teacher attendance rates), student performance on math and reading exams, and time spent on various in-class daily activities. Goodman and Turner use evidence provided by the New York City Department of Education (NYCDOE), the administrative arm of the mayor's office responsible for the city's schools. Beginning in the fall of 2007, the NYCDOE conducted a policy experiment in which 181 of the city's high-poverty schools were randomly selected to participate in a group-based teacher incentive program that provided each school the opportunity to earn schoolwide bonuses. Consistent with Figlio and Kenny's (2007) results, Goodman and Turner find no significant impact of the bonus on student achievement under the group-based incentive scheme, even in schools where the bonus was nearly identical to an individual incentive program. There was, however, a small yet significant decrease in teacher absenteeism for teachers who were eligible for the largest incentives, suggesting that these teachers may have increased their efforts in response to the incentive.

In another analysis of the New York City randomized trial, Fryer (2011) comes to essentially the same conclusions. He finds no evidence that performance-based incentives for teachers increased measured student performance, student attendance, or graduation rates. He also finds no effect of the teacher incentive scheme on student behaviors, and, unlike Goodman and Turner (2011), his analysis does not support the notion that teachers altered their behaviors in response to the incentivization. Surprisingly, he found that the teacher incentive program might have actually *decreased* student achievement, with the negative effect especially apparent in larger schools.

Initial findings of a separate randomized control trial of teacher merit pay were recently released (Springer et al. 2010) in a report issued by the research team leading the Project on Incentives in Teaching (POINT), an endeavor of the National Center on Performance Incentives at Vanderbilt University. POINT was a three-year study of middle school mathematics teachers in Nashville, Tennessee, who had volunteered for the randomized experiment. Specifically designed to test the hypothesis that substantial student performance–aligned financial incentives for teachers would increase student test scores, the POINT data suggest that merit pay did not significantly improve mathematics scores. Teachers in the experimental condition were eligible for bonuses of up to $15,000 for increasing student test scores, but they did not significantly outperform the students of teachers in the control condition.

Given that only two experimental studies examining the impact of merit pay on student performance have been conducted—and given that their results suggest that merit pay does not significantly improve student performance—more research on the nature, implementation, and effects of teacher merit pay schemes is clearly needed. Such study is particularly salient in light of the present bipartisan political favor for teacher merit pay schemes, which will be discussed in the following section. The current nonexperimental research evidence suggests that teachers may either respond to individual incentives or "free ride" when offered group-based incentives, both of which are consistent with the economic theory underpinning teacher performance pay. However, caution is necessary in interpreting the extant nonexperimental research: Figlio and Kenny (2007) remain causally agnostic about their results, and Goodman and Turner (2011) examine a highly idiosyncratic program of teacher incentives. Neither of these studies addresses policy arguments that suggest that merit pay programs are not suitable teacher compensation policies. Further inquiry into whether performance pay schemes result in goal distortion that increases test scores at the expense of a quality education would help settle the policy arguments surrounding this branch of teacher compensation reform.

## WHY CONSIDER MERIT PAY AND MUSIC EDUCATION NOW?

Given that the research on pay-for-performance teacher compensation reform in the United States is in relatively early stages, with the literature currently consisting of mostly theoretical work and little empirical evidence, one might reasonably question why it is important for music education scholarship to consider merit pay and music education now. The answer is simply that federal policymakers have forced the issue by tying teacher compensation to student performance in a key education reform policy shift of the $4.35 billion RTTT program, which is perhaps the largest federal government investment in education in American history.

Historically, schools in the United States have been funded and regulated by local communities and the states (Vinovskis 2009). To many, education seems to fall entirely under the legal purview of the states, given that the U.S. Constitution never lists regulation of the educational system as one of the enumerated powers of the federal government, and the Tenth Amendment reserves all such unspecified powers to the states (Ryan 2009). While state and local governments retain most direct regulatory control over public education, several significant policy instruments are available to the federal government to influence and regulate education. McDonnell and Elmore (1987) categorize all educational policy instruments as one of four types: (1) *mandates*, which demand nonnegotiable compliance with specific rules and actions and require the enforcement of penalties for noncompliance; (2) *inducements*, which offer significant sums of money or other investments that are conditional on the completion of specific actions and require oversight to ensure the conditions are met; (3) *capacity-building*, which offers money or other investments to promote the development of resources (human, material, or intellectual) that are expected to provide future benefits; and (4) *system-changing*, which transfers official authority to new individuals or agencies.

Whereas the No Child Left Behind Act of 2001 (NCLB) was often derided by critics as an unfunded mandate because of its strict regime of sanctions for poor-performing schools, as a reauthorization of the Elementary and Secondary Act of 1965, it is more properly classified as an inducement. Receipt of federal Title I education funds by any state or local school district was tied to these schools' acquiescence to the provisions of NCLB, and states were free to opt out entirely. This point is arguable, however, since Title I funding is essential for the continued operation of many states' education systems. RTTT, on the other hand, can be less nebulously categorized as an inducement. Under this initiative, states voluntarily enter a competitive grant process to receive a portion of the $4.35 billion fund, which is earmarked for the enactment of specific educational reforms in that state. The most interesting element of the RTTT competition is that to earn the highest possible score in the competitive process, the favored reforms—including introduction of merit pay

schemes—must have been enacted by states *prior* to their application for the funds (Goldstein 2009). Forty-one states competed for the first round of RTTT funding, and thirty-six states submitted applications for the second round. Many states have already enacted reforms induced by the program, even though only two (Delaware and Tennessee) won grants in the first round, and only nine (Massachusetts, New York, Hawaii, Florida, Rhode Island, Maryland, Georgia, North Carolina, and Ohio) plus the District of Columbia won grants in the second round (U.S. Department of Education 2010). Given the success of the RTTT thus far, merit pay schemes are poised for a national scale-up. Statewide laws enacting some form of merit pay have been enacted in the past year in Colorado, Louisiana, Oklahoma, and New York (Dillon 2010). Colorado's Ensuring Quality Instruction through Educator Effectiveness Act of 2010 (EQuITEE), for example, requires that all teachers—including teachers of nontested subjects such as music—be evaluated annually, and that at least 50 percent of the evaluation be based on longitudinal measures of student achievement.

## MERIT PAY FOR MUSIC TEACHERS: ISSUES AND CONCERNS

Since the success of the RTTT inducement has made merit pay a policy reality instead of a theoretical policy debate, arguments for and against merit pay schemes are less relevant to music educators than are the immediate concerns of implementation. Many of the policymakers and advocates who promote merit pay systems tend to focus on standardized test scores as the ultimate arbiter of teacher quality. This narrow conception of the teaching profession will likely cause challenges for the implementation of merit pay schemes, because the designers will inevitably be confronted with the diversity of teaching milieus that actually occur in the public schools. If implemented poorly, merit pay schemes could be deleterious to both the quality and number of working music teachers and the quality of the nation's music curricula. There are three main issues that music education will need to consider as merit pay systems become more widespread: (1) ensuring equitable access to the highest possible educator compensation for music teachers, (2) developing a clearer definition of music teacher quality, and (3) designing a fair and transparent system of evaluating music teacher quality to determine compensation.

Ensuring that music teachers are eligible for the same maximum compensation as teachers of other subjects—as they are under the single salary schedule—is key to maintaining the overall quality of the music teacher workforce when merit pay becomes the norm for teacher compensation. Opting out of merit pay entirely, particularly if the compensation available under merit pay schemes is greater than the compensation offered under a more traditional scheme, will not maintain the status and quality of music education.

Under the single salary schedule, the opportunity cost of becoming a teacher relative to other professions with similar educational prerequisites is high (Béteille and Loeb 2009), but this cost is equal among all types of teaching. The overall quality of the music teacher workforce could reasonably be expected to decline if the relative compensation that music teachers earn compared to nonteaching jobs falls at a greater rate than that for teachers of other subjects. That is, if music educators become ineligible for the highest possible compensation that is available to, say, a math teacher under merit pay schemes, then teaching music will become a less desirable career than teaching other subjects or pursuing a nonteaching career, because of the greater opportunity cost of becoming a music teacher. This situation already exists in Washington, DC, where music teachers rated "highly effective" are currently eligible for a pay-for-performance annual bonus that caps out at $10,000 less than that available to a similarly rated teacher of math or English (District of Columbia Public Schools 2010). Certainly, as Bergee et al. (2001) have found, we can expect that preservice music teachers will likely not choose to enter the profession because of the wages offered. However, the Bergee et al. study was conducted when the single salary schedule—under which a music teacher would earn an equal amount as a chemistry teacher of the same educational level and experience—was essentially sacrosanct. If music teachers become underpaid relative to teachers of other subjects, then future education majors who wish to *teach* more than they wish to teach *music* will likely pursue other specializations. Under such a scenario, in which music teachers will become second-class teachers paid at a lower rate, it is plausible that the available pool of music educators will shrink both in number and in quality as fewer choose to enter the profession, and that the most highly qualified teachers will exit the pool in greater numbers.

The Denver Public Schools ProComp plan (Gratz 2005; Denver Public Schools 2010) is often hailed as an exemplar of teacher compensation. Under this model, teachers are awarded incentives for meeting collaboratively determined "student growth objectives," attaining graduate degrees, attaining National Board Certification, and working in hard-to-staff schools or hard-to-staff subjects. Music teachers are eligible for all of the numerous incentives available to other teachers, except for the single $2,400 "Exceeds Expectations" incentive tied to test scores, which is reserved for teachers of tested subjects in tested grades. While the Denver plan represents a very strong compromise, a clear "second class" of compensation still exists for those teachers who do not teach fourth- through twelfth-grade math or English, who are the only teachers in Denver eligible for the highest possible compensation under ProComp.

The second issue for music teaching to consider as merit pay reforms are enacted is our own definition of teacher quality and how we might link teacher quality to student achievement. Merit pay schemes are founded on the assumptions that (1) the best teachers elicit the greatest student learning and

student achievement growth possible, and (2) this achievement growth is validly, accurately, and reliably measurable. The extant music education research on teacher quality or teacher effectiveness is somewhat inadequate as a basis for designing a compensation scheme that links teacher pay to student achievement. This inadequacy stems from the fact that most research on music teacher quality does not use measures of student achievement in music as the tested outcome (Taebel 1992). Instead, these studies tend to focus on the ratings made by expert (preservice or in-service music teachers) or novice (secondary student) judges who view a recorded teaching episode and focus their attention on teacher behaviors, rather than student achievement. In a comprehensive review of this literature, Duke writes,

> Implicit in the use of the word "effective" to describe teaching is the notion that the effect of interest is a positive change in some aspect(s) of student behavior—what students know or are able to do. Yet, this central purpose of teaching is grossly underrepresented in the dependent measures employed in [music education] research. (1999, 16–17)

In other words, for the majority of the research that both Duke (1999) and Taebel (1992) reviewed, the question "Are you an effective teacher?" was not tantamount to the question "Are your students learning?" Because of this disconnect, schools rarely employ the research literature on music teacher quality when they evaluate music teachers in the field. Maranzano (2000) suggests that partly as the result of the lack of solid research on music teacher quality and evaluation, the evaluation practices in wide use today are not successful in identifying music teachers' content-specific strengths and weaknesses. As such, current teacher evaluation practices are likely to be unsuccessful if used for compensation-related evaluation.

The situation linking student achievement measures to teacher effectiveness has not improved in the years since Duke's review. A study by Hamann et al. (2000) suggests that university music students asked to rate teaching effectiveness will rate teaching highly if the delivery of the instruction is engaging, even if the content of the lesson is poor. In this study, music teachers with engaging delivery styles but lesson content that veered off the lesson plan (and sometimes off of educational content entirely) were rated more highly than teachers with poor delivery and quality instructional content. Madsen (2003) has shown that even inaccurate instruction can receive high ratings from student raters, if the content is delivered in an engaging manner. More recently, Johnson, Price, and Schroeder (2010) sought to determine whether teaching effectiveness could be rated *irrespective* of the proficiency of the choral ensemble—that is, they aimed to test whether previous research findings about engaging delivery would hold when a perceptible disconnect existed between the quality of the teaching and the student achievement demonstrated in the lesson. The researchers conclude that their raters could discern a teacher's skill independently of the quality of the music made by the students. They suggest that "this finding is important because it unmistakably shows that participants could demonstrate the ability to look at one aspect of a teaching even while ignoring several other distractions, such as student performance" (14). It is unfortunate that the line of research on teacher quality in music has become conceptually disconnected from the idea that Duke (1999) puts forward—that effective teaching is by necessity linked with better student achievement—given that student achievement will be the de facto measure used for evaluating and compensating music teachers in states adopting the RTTT-induced reforms.

The profession's definition of music teacher quality, then, must be aligned with student achievement in music to provide a solid research basis upon which adequate teacher evaluation methods can be built. In his call for more research that examines student achievement as the primary indicator of teaching effectiveness, Duke (1999) sidesteps the issue of a lack of consensus regarding what desirable student outcomes in music education should be. In an article describing the work of the task force charged with determining whether revisions should be made to the National Standards in music, Lehman notes that the most striking feature of the results of a music educator survey that the task force commissioned to seek proposals for changes to the standards was the "the spectacular lack of unanimity among respondents about what changes should be made in the standards" (2008, 32). During the mid-1990s, the National Standards were developed during a long process involving extensive field-wide consultation and careful deliberation from a distinguished panel of music educators. For the purposes of developing a merit pay scheme that links teacher effectiveness with student outcomes, the National Standards seem to be the most thoroughly vetted framework our profession has, despite the lack of unanimity that may exist over their current form or content.

The final issue that music educators need to consider in light of the emergence of merit pay schemes is the design and development of a fair and transparent system of evaluating music teacher effectiveness that is suitable for determining compensation. Neal (2009) offers a useful suggestion for designing merit pay schemes: the crafters of such a system must deliberately designate priorities. Merit pay proposals must first "clearly delineate the types of achievement that the system is intended to foster" and then ensure that "the mapping between the policy priorities that define [the] incentive system for educators and the procedures used to create performance rankings . . . for teachers should be clear and precise" (155). This identification is important for music educators in particular, because music and music study are different from the study of, say, math and English in a number of important ways that must be recognized in an equitable compensation scheme. Neal's suggestions can help designers develop a well-designed pay-for-performance plan and, in doing so, would likely heed Rothstein's warning to avoid "the folly of rewarding A while hoping for B" (2009b, 98).

As pay-for-performance schemes induced by RTTT grants are implemented in the states, music teachers and state music educators' associations must take an active role in developing these systems to avoid being relegated to a second-class status. There are many likely allies in the design and implementation of a workable pay-for-performance scheme for music educators. The pro–merit pay Center for Educator Compensation Reform within the U.S. Department of Education (Prince et al. 2009) estimates that nearly 69 percent of the teacher workforce currently teaches a nontested subject or a nontested grade and proffers suggestions for designing fair and transparent teacher evaluation systems for merit pay programs that incorporate teachers of nontested subjects. Prince et al. suggest four options for policymakers interested in designing merit pay schemes that fairly compensate teachers of subjects such as music. These four options are: (1) restrict the merit pay portion of compensation to schoolwide performance bonuses for all teachers; (2) provide teachers of nontested subjects with the opportunity to earn some individual performance incentives, but restrict other incentives to tested-subject teachers only; (3) create new standardized tests for all subject areas, including music; and (4) base merit pay for nontested teachers "exclusively on non-test measures, such as observed evaluations of classroom performance, acquisition of additional knowledge and skills . . . and other non-standardized test measures [that] may be displayed in a thoughtful portfolio" (17). The first option should be rejected on the basis of the empirical evidence reviewed earlier (Figlio and Kenny 2007; Fryer 2011; Goodman and Turner 2011). The second option should be rejected on the premise that music teachers need to remain eligible for the maximum compensation that is offered to nonmusic teachers to maintain the quality of the workforce. The third option is likely to receive the most resistance from music educators, who do not wish to resort to pencil-and-paper standardized tests in their subject area. The fourth option, however, has great promise.

In the remainder of this article, I will propose the basic framework of one possible merit pay system that could fairly compensate teachers of music and follows the suggestions of Neal (2009), Prince et al. (2009), Taebel (1992), and other scholars who have thoughtfully considered merit pay plans or the evaluation of music teachers. The framework proposed here is not designed as a turnkey solution that could be effortlessly implemented in all contexts, but it will hopefully serve as an important building block to inform statewide music educators associations and others seeking to become a voice for music teachers in the design and scale-up of merit pay plans.

## FRAMEWORK OF A POSSIBLE MUSIC EDUCATION PAY-FOR-PERFORMANCE PLAN

Constructing a fair merit pay program—even one that is only designed to reward the 31 percent of teachers who work with students in tested grades or tested subjects—is not a simple task. Teaching is complex and multifaceted work, and student achievement in any of the disciplines is often highly influenced by sociocultural factors well beyond the purview of schools and classrooms (Berliner and Biddle 1996). Within music teaching itself, "the diversity of tasks among general, choral, and instrumental music teachers makes consensus on evaluative criteria difficult" (Taebel 1992, 311). Although not perfect, the Denver ProComp plan serves as a useful model, as do recommendations from Neal (2009), Prince et al. (2009), and other scholars, all of which have been considered in developing the framework proposed here.

Neal (2009) recommends that at the outset of designing a merit pay plan, policymakers should determine the priorities they wish to incent. Achieving consensus on a list of priorities for music teaching and learning has been difficult for music teachers and music education scholars (Taebel 1992). Even though the National Standards in music remain somewhat controversial within the profession some sixteen years after they were first developed (Lehman 2008), I believe that they still represent the most thoroughly vetted and comprehensive list of aims for student learning that we have. Thus, following Neal's (2009) suggestion, we can clearly identify the priority of this merit pay system for music teachers: advancing students' level of understanding and achievement in the curricular areas represented by the National Standards that are applicable to their teaching assignment. Since the National Standards are comprehensive in nature, even teachers with relatively narrow teaching assignments (such as secondary choral music or elementary general music) can find appropriate goals for student growth within the nine content standards. Thus, in this proposed pay-for-performance framework, the clear priority is improving student achievement in one or more areas of the National Standards.

Taebel asserts that "the complex nature of [music] teaching makes it impossible to evaluate a teacher's performance with a single evaluation instrument" (1992, 316). He recommends the use of multiple lines of evidence in evaluating the work of music teachers. In considering a merit pay proposal that is fair to music teachers, the Denver ProComp model (Denver Public Schools 2010) in which teachers are eligible for many possible incentives that are each paid at a varying rate is useful. The Denver ProComp model was briefly outlined earlier; a complete review of its features here shows how it can serve as a starting point for a merit pay system that is workable for music educators.

Under ProComp, in 2010–11, all participating teachers became eligible for base salary increments or bonus incentives linked to (1) completion of professional development courses, (2) attainment of an advanced degree or a National Board certification, (3) the principal's annual evaluations, (4) service in a hard-to-staff or hard-to-serve school, (5) achievement of teacher-designed annual goals for student learning, (6) service in a school designated as a districtwide top performer, or (7) service in a school designated as a "high

growth" school. Teachers in tested grades or tested subjects are eligible for an additional increment to their base salary if their students exceed expectations on the state test. The initial base salary for a Denver ProComp teacher is determined using a traditional salary schedule that accounts for years of experience and educational attainment. Future years' raises are calculated entirely as a function of the "base adding" incentives the teacher earns each year, and they accumulate across a teacher's career; "bonuses," as the name implies, are one-time awards that are not paid again the following year unless the objective is met again. The Denver model could be useful in developing other systems, provided that initial base salary amounts remain competitive and base-adding salary incentives are not unduly difficult to attain. Again, the primary issue with the Denver model for music teachers is that they remain ineligible for the highest level of compensation.

A small adjustment to the Denver ProComp model could reinstate parity for music teachers. Merit pay plans use standardized tests essentially as external audits of student achievement. It would not be infeasible to replicate this basic functionality—an external audit of student achievement—in music programs. Pencil-and-paper tests are likely not the solution: music educators generally do not view these exams to be authentic assessments of music learning. Few, if any, of the National Standards for Music Education could be adequately assessed in this manner in a way that would be favorably viewed by the majority of music teachers. However, for music teachers in performance areas, it would be relatively easy to audio record students at the start of the school year, audio record them at the end of the school year, and submit the audio recording to an external auditor for assessment. Teachers of nonperformance music areas could likewise create some kind of pretest/posttest portfolio record of student work and submit it for an external evaluation. Creating an infrastructure for this kind of evaluation need not be exceedingly difficult; in fact, precedent for precisely this kind of external evaluation of music teachers' work already exists. Music teachers seeking National Board certification currently submit four portfolio entries to the NBPTS, each of which consists of a video demonstration that features the teacher's delivery of instruction and students' musical responses (NBPTS 2010). Additionally, teachers are required to submit written commentary reflecting on the submitted teaching episodes. Using the Educational Testing Service (ETS) as a subcontractor, NBPTS hires experienced music teachers to evaluate the submitted portfolio entries.[3] The subtle difference between NBPTS evaluation and my proposal is that the teacher's delivery of instruction would not be the focus of the external audit; instead, that delivery would remain under the purview of the school principal or music supervisor. This external audit would focus only on student achievement, maintaining its parallel to the pencil-and-paper test in the merit pay scheme.

Under the external audit scenario, music teachers or their administrators could be free to choose the subset of the National Standards that most closely matches the expectations of their teaching assignment. The submission of performance recordings or portfolios of general music student work could be accompanied by written commentary from the teacher, as NBPTS requires, although this requirement could be left to the discretion of the school district or the auditing organization. The auditing organization would need to maintain transparency by disseminating the rubrics and the "anchor" recording and portfolio samples used to train auditors. School districts could be free to choose a "growth" model or an absolute achievement standard by which they would judge their teachers; this option should mirror the method chosen to evaluate teachers of tested subjects. In his discussion of how large-scale assessments in music could work, Shuler (2008) suggests that technological advances in digital recording make the capturing, submission, and scoring of audio recordings relatively simple: once recorded digitally, audio could be uploaded and even remotely scored by trained evaluators entirely over the Internet.

Interestingly, the external audit scenario may provide a policy lever that would allow scholars who have sought to expand the comprehensiveness of music education beyond the traditional bands, choirs, and orchestras (i.e., Kratus 2007; Williams 2007) to see their ideas become viable in the public schools. If schools chose to hold their music teachers accountable for more than, say, standards 1, 2, and 5 (singing, playing, and reading notation), then a meaningful curricular readjustment might occur. The use of merit pay as a policy lever to influence curricular change in music becomes a particularly intriguing possibility when one considers that the writers of the original National Standards for Music (Consortium of National Arts Education Associations 1994) likely hoped that the standards themselves would serve as a policy lever to broaden the scope of the nation's music education curriculum. Ultimately, the music National Standards failed to significantly alter the landscape of music education in the schools because they lacked enforcement of any kind (being dubbed "voluntary," as were the standards released in all other subject areas) and because of the charged political rhetoric that surrounded the entire effort to create National Standards in all subject areas (Ravitch 2010). The arts standards in particular served as a political whipping post for those opposed to National Standards from all sides of the political spectrum. Upon the release of the National Standards for the Arts, Albert Shanker, then president of the American Federation of Teachers, famously derided them as "ridiculous" and "a wish list" (quoted in Mark 1996, 109). Using the external audit scenario as a policy lever to alter the content of the music curriculum, however, would require significant buy-in from those schools choosing to use the services of the auditor and would remove the ability, suggested earlier, to tailor the audit process to the teaching assignment as it exists. The work of scholars who wish to broaden the scope of music in the schools would thus need to be directed toward convincing the auditors and the schools that such a shift in school music is truly warranted.

## NEXT STEPS IN RESEARCH AND PRACTICE

As policy, teacher compensation reform seems likely to enjoy its current bipartisan support for some time to come. Although research outside of music education has suggested that merit pay schemes fail in improving student performance by aligning it with teacher incentives, merit pay plans have become the operating procedure in a large number of U.S. school districts. As a result, the profession will need to develop its own body of systematic knowledge on the effects of merit pay schemes on music educators and music education. Qualitative research in music education ought to interview music educators working in some of the more high-profile merit pay school systems and states (e.g., Washington, DC, Charlotte-Mecklenburg, N.C., districts in Colorado under the new EQuiTEE Act). This kind of in-depth qualitative work can illuminate perceptions of merit pay as beneficial or deleterious to working conditions, job satisfaction, teacher migration or attrition, and a host of other elements of the music teaching profession. Understanding whether merit pay has negative consequences for music teacher working conditions is an important first step in understanding whether merit pay causes consequential shifts in the music teacher labor market. As Duke (1999) has contended, quantitative research should seek to develop measures of teaching effectiveness that are more linked to student learning than to observable teacher characteristics and behaviors, so that frameworks like the one proposed here might have a more solid research-based grounding.

Music educators in schools facing the adoption of merit pay must work locally to ensure that they are not simply ignored in the process. In some states, such efforts might encourage music teachers to become more active in their collective bargaining units to ensure that local contracts preserve compensation equity regardless of teacher compensation method. In states without strong collective bargaining frameworks, music teachers must find other ways to be represented in the process, either through district committee work or the encouragement of political action by music parents (Elpus 2008). The framework proposed here might serve as a starting point in the discussion about music education's production of student learning that is measurable, valid, and important.

Ultimately, many substantial logistical details will need to be considered to create an external auditing system that is transparent, fair, and effective at determining the quality of student learning in music. In the current policy and political climate clamoring for more "teacher accountability," the profession of music education can no longer avoid the issue of measuring student musical learning in meaningful ways. The winning states of both the first and second rounds of funding of RTTT have been announced, and in many of these states, legislation that will fundamentally alter the traditional methods of teacher evaluation have been passed or are under debate. Even in states that were not successful in the initial iterations of RTTT, reformers' calls for new methods of teacher evaluation are being heeded. If music educators cannot articulate a method of evaluation that meets the new standards, we run the risk of forgoing the salary equity with teachers of other subjects that we have historically enjoyed—an outcome that will likely exacerbate music teacher shortages and inequities of access to music education.

## NOTES

1. For 2010–11, K–12 public school music teachers are eligible for Perkins loan cancellation in Alabama, parts of Arizona, Connecticut, Delaware, the District of Columbia, Georgia, Idaho, Iowa, Missouri, Montana, Nevada, New York City, North Dakota, South Carolina, South Dakota, Vermont, Wisconsin, and the U.S Virgin Islands (U.S. Department of Education 2009).

2. Interestingly, Glewwe, Ilias, and Kremer (2003) and Lavy (2002) find conflicting results regarding the integrity of the incentive system. Lavy finds no evidence of teachers "gaming the system," while Glewwe, Ilias, and Kremer conclude that short-term gains on the test used to measure teacher performance were not associated with greater long-term student achievement.

3. I know this process from experience, having once been solicited by the ETS to serve as an NBPTS music teaching evaluator shortly after I earned my master's degree in music education while I was teaching at a school near the ETS main campus in Princeton, New Jersey.

## REFERENCES

Adams, S. J., and J. S. Heywood. 2009. Performance pay in the U.S. private sector: Concepts, measurement, and trends. In *Teachers, performance pay, and accountability: What education should learn from other sectors*, ed. S. J. Adams, J. S. Heywood, and R. Rothstein. Washington, DC: Economic Policy Institute.

Bergee, M. J., D. D. Coffman, S. M. Demorest, J. T. Humphreys, and L. P. Thornton. 2001. *Influences on collegiate students' decision to become a music educator*. MENC: The National Association for Music Education 2001. http://www.menc.org/resources/view/influences-on-collegiate-students-decision-to-become-a-music-educator.

Berliner, D. C., and B. J. Biddle. 1996. *The manufactured crisis: Myths, fraud, and the attack on America's public schools*. New York: Perseus.

Béteille, T., and S. Loeb. 2009. Teacher quality and teacher labor markets. In *Handbook of education policy research*, ed. G. Sykes, B. Schneider, and D. N. Plank, 596–612. New York: Routledge.

Buck, S., and J. P. Greene. 2011. Blocked, diluted, and co-opted: Interest groups wage war against merit pay. *Education Next* 11 (2): 26–31.

Campbell, D. T. 1976. Assessing the impact of planned social change. http://www.eric.ed.gov/PDFS/ED303512.pdf.

Consortium of National Arts Education Associations. 1994. *National standards for arts education: What every young American should know and be able to do in the arts*. Reston, VA: MENC: The National Association for Music Education.

Dee, T. S., and B. J. Keys. 2004. Does merit pay reward good teachers? Evidence from a randomized experiment. *Journal of Policy Analysis and Management* 23 (3): 471–88.

Denver Public Schools. 2010. *Welcome to teacher ProComp*. http://denverprocomp.dpsk12.org/.

Dillon, S. 2010. States create flood of education bills. *New York Times*, May 31, A14.

District of Columbia Public Schools. 2010. *IMPACTplus guidebook*. Washington, DC: Author. http://www.dc.gov/DCPS/Files/downloads/TEACHING%20&%20LEARNING/IMPACT/IMPACTplus/DCPS-IMPACTplus-guidebook-Sept-2010.pdf.

Duke, R. 1999. Measures of instructional effectiveness in music research. *Bulletin of the Council for Research in Music Education* 143:1–48.

Elpus, K. 2008. Organizing your parents for effective advocacy. *Music Educators Journal* 95 (2): 56–61.

Figlio, D. N., and L. W. Kenny. 2007. Individual teacher incentives and student performance. *Journal of Public Economics* 91 (5/6): 901–14.

Fryer, R. G. 2011. Teacher incentives and student achievement: Evidence from New York City public schools. NBER Working Paper 16850. Cambridge, MA: National Bureau of Economic Research.

Glewwe, P., N. Ilias, and M. Kremer. 2003. Teacher incentives. NBER Working Paper 9671. Cambridge, MA: National Bureau of Economic Research.

Goldstein, D. 2009. The selling of school reform. *The Nation*, June 15, 18–22.

Goodman, S., and L. Turner. 2011. Does whole-school performance pay improve student learning? Evidence from the New York City schools. *Education Next* 11 (2): 66–71.

Gratz, D. G. 2005. Lessons from Denver: The pay for performance pilot. *Phi Delta Kappan* 86 (8): 568.

Hamann, D. L., D. S. Baker, P. A. McAllister, and W. I. Bauer. 2000. Factors affecting university music students' perceptions of lesson quality and teaching effectiveness. *Journal of Research in Music Education* 48 (2): 102–13.

Hanushek, E. A. 2003. The failure of input-based schooling policies. *Economic Journal* 113 (485): 64–98.

Jacob, B. A., and S. D. Levitt. 2003. Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *Quarterly Journal of Economics* 118 (3): 843–77.

Johnson, C. M., H. E. Price, and L. K. Schroeder. 2010. Teaching evaluations and comments of pre-service music teachers regarding expert and novice choral conductors. *International Journal of Music Education* 27 (1): 7–18.

Jones, P. M. 2009. Hard and soft policies in music education: Building the capacity of teachers to understand, study, and influence them. *Arts Education Policy Review* 110 (4): 27–32.

Kos, R. P., Jr. 2010. Developing capacity for change: A policy analysis for the music education profession. *Arts Education Policy Review* 111 (3): 97–104.

Kratus, J. 2007. Music education at the tipping point. *Music Educators Journal* 94 (2): 42–48.

Lavy, V. 2002. Evaluating the effect of teachers' group performance incentives on pupil achievement. *Journal of Political Economy* 110 (6): 1286–317.

Lazear, E. P. 2003. Teacher incentives. *Swedish Economic Policy Review* 10:179–214.

Lehman, P. 2008. A vision for the future: Looking at the standards. *Music Educators Journal* 94 (4): 28–32.

Madsen, K. 2003. The effect of accuracy of instruction, teacher delivery, and student attentiveness on musicians' evaluation of teacher effectiveness. *Journal of Research in Music Education* 53 (1): 38–50.

Maranzano, C. 2000. Music teacher performance evaluation: A call for more inclusive models. *Journal of Personnel Evaluation in Education* 13 (3): 267–74.

Mark, M. L. 1996. *Contemporary music education*. 3rd ed. New York: Schirmer.

McDonnell, L. M., and R. F. Elmore. 1987. Getting the job done: Alternative policy instruments. *Educational Evaluation and Policy Analysis* 9 (2): 133–52.

National Board for Professional Teaching Standards. 2001. *NBPTS music standards*. Arlington, VA: Author.

———. 2010. *Guide to national board certification*. Arlington, VA: Author.

Neal, D. 2009. Designing incentive systems for schools. In *Performance incentives: Their growing impact on American K–12 education*, ed. M. G. Springer, 149–70. Washington, DC: Brookings Institution Press.

Podgursky, M. 2009. A market-based perspective on teacher compensation reform. In *Performance incentives: Their growing impact on American K-12 education*, ed. M. G. Springer, 67–86. Washington, DC: Brookings Institution Press.

Podgursky, M., and M. G. Springer. 2007. Teacher performance pay: A review. *Journal of Policy Analysis and Management* 26 (4): 909–50.

Prince, C. D., P. J. Schuermann, J. W. Guthrie, P. J. Witham, A. T. Milanowski, and C. A. Thorn. 2009. *The other 69 percent: Fairly rewarding the performance of teachers of non-tested subjects and grades*. Washington, DC: Center for Educator Compensation Reform. http://cecr.ed.gov/guides/other69Percent.pdf.

Ravitch, D. 2010. *The death and life of the great American school system: How testing and choice are undermining education*. New York: Basic Books.

Rockoff, J. E., B. Jacob, T. J. Kane, and D. O. Staiger. 2008. Can you recognize an effective teacher when you recruit one? NBER Working Paper 14485. Cambridge, MA: National Bureau of Economic Research.

Rothstein, R. 2009a. The influence of scholarship and experience in other fields on teacher compensation reform. In *Performance incentives: Their growing impact on K–12 education*, ed. M. G. Springer, 87–110. Washington, DC: Brookings Institution Press.

Rothstein, R. 2009b. The perils of quantitative performance accountability. In *Teachers, performance pay, and accountability: What education should learn from other sectors*, ed. S. J. Adams, J. S. Heywood, and R. Rothstein, 69–107. Washington, DC: Economic Policy Institute.

Ryan, J. E. 2009. A legal perspective on teacher compensation reform. In *Performance incentives: Their growing impact on American K–12 education*, ed. M. G. Springer, 43–66. Washington, DC: Brookings Institution Press.

Scott, J. C. 1998. *Seeing like a state: How certain schemes to improve the human condition have failed*. New Haven, CT: Yale University Press.

Shuler, S. C. 2008. Large-scale assessment of music performance: Some whys and hows for today's data-driven educational environment. In *Assessment in music education: Integrating curriculum, theory, and practice*, ed. T. Brophy, 123–37. Chicago: GIA.

Springer, M. G. 2009. Rethinking teacher compensation policies: Why now, why again? In *Performance incentives: Their growing impact on American K–12 education*, ed. M. G. Springer, 1–22. Washington, DC: Brookings Institution Press.

Springer, M. G., D. Ballou, L. Hamilton, V.-N. Le, J. R. Lockwood, D. F. McCaffrey, M. Pepper, and B. M. Stecher. 2010. *Teacher pay for performance: Experimental evidence from the Project on Incentives in Teaching*. Nashville, TN: National Center on Performance Incentives at Vanderbilt University.

Taebel, D. K. 1992. The evaluation of music teachers and teaching. In *Handbook of research on music teaching and learning*, ed. R. Colwell, 310–27. New York: Schirmer.

U.S. Department of Education. 2009. *Teacher shortage areas: Nationwide listing*. Washington, DC: Author.

———. 2010. *Nine states and the District of Columbia win second round Race to the Top Grants*. http://www.ed.gov/news/press-releases/nine-states-and-district-columbia-win-second-round-race-top-grants.

Vinovskis, M. 2009. *From A Nation at Risk to No Child Left Behind: National education goals and the creation of federal education policy*. New York: Teachers College Press.

Weisberg, D., S. Sexton, J. Mulhern, and D. Keeling. 2009. The widget effect: Our national failure to acknowledge and account for differences in teacher effectiveness. Brooklyn, NY: The New Teacher Project. http://widgeteffect.org.

Williams, D. A. 2007. What are music educators doing and how well are we doing it? *Music Educators Journal* 94 (1): 18–23.