

## **On the Origin and Political Significance of Test-Based Teacher Evaluation and Compensation**

**Mark J. Garrison**  
*D'Youville College*

### ***Introduction***

Education “reform”<sup>1</sup> is unfolding at an unprecedented rate, with little public input. The model of corporate school reform—including charter chains, increased testing controlled by private publishing houses, the elimination of public governance mechanisms, and the emasculation of teacher unions—is pushed by some of the world’s wealthiest and most powerful champions of “change,” what some observers have called the “billionaire boys club” (Ravitch, 2010). Thus, a cabal of CEOs, corporate-interest front groups, venture philanthropists, and hedge fund managers heavily influence the nature, character and focus of education reform (Emery & Ohanian, 2004; Hass, 2009; Kovacs, 2011; Saltman, 2010). What was not possible ten years ago is now sweeping the nation: the elimination of tenure and establishment of a standardized test-based evaluation and compensation scheme for teachers (e.g., the *Student Success Act* signed into law by Florida Governor Scott in March 2011). In a short time, the use of student test scores to compensate and/or evaluate teachers has gone from a fantasy of neoliberal ideologues to policy reality for teachers in at least 20 states and the District of Columbia (California, Colorado, Connecticut, Delaware, Florida, Georgia, Hawaii, Illinois, Louisiana, Maryland, Massachusetts, Michigan, New York, North Carolina, Ohio, Rhode Island, Tennessee, Texas, West Virginia and Wisconsin). (See reports from the National Council on Teacher Quality (<http://nctq.org>) for uncritical yet instructive reviews of state and district teacher evaluation plans that accompanied their *Race to the Top* applications.)

The most discussed factor driving this transformation is the nearly \$5 billion in discretionary funding provided to the US Department of Education (USDOE) by the *American Recovery and Reinvestment Act of 2009* (ARRA), known as *Race to the Top* (RTT). This program uses taxpayer funds to bribe states into compliance with increasingly unpopular USDOE reforms, including the use of test scores to evaluate teachers. Obama’s “blueprint” for education, the administration’s model for the reauthorization of the *Elementary and Secondary Education Act*, likewise includes provisions for increased student testing and the use of student

test scores for the evaluation of teachers based on student test scores (US Department of Education, 2010).

This article offers an analysis of these historic changes now taking place by exploring the origin and political significance of test-based teacher evaluation and compensation (hereafter noted as TBTEC for purposes of brevity). The push for TBTEC is connected to broad changes in local and national governing arrangements as they relate to official claims that there is simply not enough money in the public treasury to support public education as in the past. The connection between demands for the radical restructuring of public education in general and TBTEC in particular, and the demand of the wealthy elite to cut spending on public education (and nearly every social program), provides an important starting point for the study of the origins of and political significance of TBTEC systems from a historical, economic and political vantage point. To this end, the paper asks: what are the origins and political significance of current TBTEC schemes and what does this analysis suggest for effective opposition to current education “reform” efforts?

The paper argues that the impetus for TBTEC is the drive of a faction of the super rich to cheapen education, with two distinct aims: to reduce the total financial expenditure on public schooling (especially for those students who are the object of “achievement gap” rhetoric) and to lower the quality of education by converting public schooling (again especially for those schools attended by students who are the object of “achievement gap” rhetoric) into a guaranteed revenue/profit generator for venture capital and other corporate and banking interests. The effort to control teachers through TBTEC should also be connected to what appears as the social control agenda of corporate school reform, the creation of compliant worker, non-citizens (e.g., Horn, 2011). Thus TBTEC is viewed as part of the imposition of a corporate governance structure, in place of one organized under the principal of popular election at the local and state level. Finally, it is suggested that the 19th century idea of public education as the “great equalizer” has failed as it systematically fails to address the social, economic and political roots of equality. The present conditions call for researchers and educators to rethink the content of public schooling and its role in winning the battle for democracy.

### ***Method and Definitions***

A summation of both research and experience predicts that test-based teacher compensation and evaluation will do little to improve the quality of education, and possibly do much to further erode it (e.g., Baker, Barton, Darling-Hammon, Haertel, Ladd, Linn, et al., 2010; Coltham, 1972; Fryer, 2011; Gratz, 2009; Helper, Kleiner & Wang, 2010; Johnson, 1984; Kupermintz, 2003; Martins, 2010; Murnane & Cohen, 1986; Perry, Engders & Jun, 2009; Rapple, 1994; Ravitch, 2010; Zhang, 2002)<sup>2</sup>. Taking this summation into account, the tendency of reformers to deem all evidence contrary to their cause “irrelevant” (e.g., Braun, 2010) and insist that the issue is to “believe” in the “reforms” (Garrison, 2011a) suggest that in order to examine the origin and political significance of TBTEC, analyses must go beyond empirical research regarding the efficacy of what is often called merit or performance pay. Put differently, the origin of TBTEC will not be found in the reform rhetoric of “closing achievement gaps” and improving education since there is so little evidence that such a policy will contribute to either of those oft-stated goals.

### ***Understanding and Analyzing Origin***

It is important to understand that origin refers, first and foremost, to the conditions that give rise to a thing or phenomenon. To study something’s origin is not the same as asking what “caused” something in the sense of a laboratory experiment; a better synonym for origin is *impetus*. To explore origin means to ask: what conditions and contexts are required for something to occur? That similar conditions may yield different outcomes means that no simple cause and effect model is assumed for such an analysis. Yet, by examining the conditions that existed during previous episodes of the use of exams to evaluate and compensate teachers, the most important social forces driving present efforts might be identified (Garrison, 2009).

To identify points of origin for the drive to impose TBTEC, the extensive yet rarely discussed 30-plus year performance pay schemes imposed on English and Irish schoolteachers in the latter part of the 19th century are analyzed. This “payment-by-results” scheme, as it was called, was the first modern, large-scale use of examinations as a basis for teacher evaluation and compensation. By determining the impetus for the adoption of “payment-by-results” in the 19th century, important factors giving rise to TBTEC in the present can also be identified.

### ***A Heuristic for Analyzing the Political Significance of Reform***

Something has “political significance” if it is connected to changes in governing arrangements. “Arrangements” refers not to party politics per se, but to the nature of political power, how and by whom it is wielded, and the justification of the right to that power (arrangements may be shaped more by one party or faction than another). The political significance of TBTEC is determined by examining if and how it alters power relations among teachers, and between teachers, students, administrators, and parents; it also asks how TBTEC alters power relations between teacher organizations and other entities. Finally, TBTEC should be analyzed to see if and how it shifts decision-making from public to private entities, and thus shifting to whom and for what purpose TBTEC systems are accountable.

These questions regarding the political significance of TBTEC are premised on the idea that standards (such as those adopted for the evaluation and compensation of teachers) should be studied as “weapons in the quest for power, as a means by which to give material expression to a philosophy or aim, and as embodiments of the social values of a culture or class.” Efforts to changes standards are thus explored as means to exact changes in political power, philosophy and social values (Garrison, 2009, p. 4).

### ***The Difference Between Merit Pay and Performance Pay***

It should be clear to the reader that neither the language of merit pay or performance pay has been adopted here. The reason for this is simple: there is little evidence that TBTEC has anything to do with teacher value (merit) or captures the effects of teachers’ work with students (performance). To use these terms as they appear in the reform discourse would serve to further dis-/mis-inform an already problematic discussion. Some discussion of the distinction between merit pay and performance pay, however, is warranted and useful.

To compensate an employee based on performance means to pay them based on some measure of the outcome (or output) of their efforts. These forms of compensation are historically and presently linked to ideas about motivation. “The concept of using incentives as well as coercion to induce and influence performance reaches back into our earliest recorded history” (Peach & Wren, 1992, p. 5). Practices that resemble what we today understand as performance pay are found in the Code of Hammurabi (circa 18th century B.C.), the “sixth and greatest king of the first dynasty of Babylon” (p. 7) (present-day Iraq). Piece rates are a good example of such

a system. In ancient Babylon, according to Peach and Wren, weavers of cloth were paid “in food, the quantity granted depending upon the spinner or weavers' output” (p. 7).

Thus, a central feature of performance pay is the rejection of time as the basis for compensation. With compensation systems based on time, the employer is provided temporary control over the use of the abilities of the employee. Performance pay, on the other hand, focuses on the output of the employee, not the time spent on tasks related to the employee's job description. This is key as TBTEC schemes are being accompanied by demands that teachers work longer hours, while simultaneously forgoing seniority (another measure of time tied to compensation).

Recent analyses suggest, however, that piece rate systems (individual incentive pay based on a measure of output), which TBTEC systems model, are consistent with industrial models of production where worker tasks are simple and easy to observe. Modern manufacturing methods, on the other hand, are inconsistent with piece rate systems as many worker tasks are neither simple nor easy to observe (Helper, Kleiner & Wang, 2010). Thus the justification that TBTEC should be adopted because it is more consistent with “professionalism” and not based on the “outdated industrial model” where “all workers are treated the same” is patently false (e.g., Finne, 2010; Weisberg, Sexton, Mulhern, & Keeling, 2009). Note that while group incentives are often positioned as a better alternative to individual incentive plans, recent research based on experiments in New York City show no positive effect of group-incentive pay based on student test scores, and some evidence that group-incentive pay had a *negative* effect on test scores (Fryer, 2011).

Despite their conflation in current policy talk, performance pay is quite different from merit pay, which involves some subjective judgment of a supervisor. Merit pay may be awarded for a “job well done” as judged by the boss, but such designations are not typically attached to some “measurable” output. While some current TBTEC schemes award “bonuses” (or “merit pay”) to teachers whose students show large gains on standardized tests, these bonuses are best understood as forms of performance pay and not merit pay, as the measure is still essentially piece-rate, akin to a bonus for a salesperson who crosses a threshold of sales in a week (e.g., Sarrio, 2011).

Importantly, despite their public support for it, most businesses do not adopt performance (as opposed to merit) pay, except in situations where the output is easy to observe, as in piece-

rate contexts. Most business that claim to use performance pay in fact use combination of subjective and objective measures, relying heavily on supervisors ratings for bonuses (Gratz, 2009). Yet, current proposals typically recommend that at least 50 percent of a teacher's evaluation be based on test scores (and hence a decision to retain or fire the teacher), indicating a move to a performance pay structure, since the judgment of direct supervisors is diminished compared to past practice.

### ***The 19th Century Experiment with "Payment-by-Results"***

"Payment-by-results," as it was called, constituted a more than 30-year experiment within the United Kingdom (England and Ireland) following the Revised Code of 1862 (ending in the mid 1890s in England and 1920s in Ireland, with the political split of Ireland in 1922). While earlier plans for performance pay in elementary education were attempted (Burton, 1979; Rapple, 1994, p. 3), the Revised Code adopted the heretofore most significant and extensive application of a TBTEC-like system (see Coltham, 1972; Dent, 1963; Madaus, Ryan, Kellighan & Airasian, 1987; Rapple, 1994; for similar efforts in Australia see Griffiths, 1987). Throughout its existence the program maintained the following characteristics: allocation of funds to schools and teachers based on attendance and attainment of students, as determined by Her Majesty's Inspectors, with some provisions to provide monetary rewards directly to students. While teachers were not compensated on the basis of what we today would understand as the results of standardized tests or "growth models," inspectors were provided scripted examination protocol, the results of which determined a majority of teacher pay. In 1862 the Committee of Council on Education "set out very specific instructions to the inspectors concerning the administration of the annual examination. Inspectors were advised that the test in the three Rs 'of individual children according to a certain standard must always be, to a considerable extent, mechanical'" (Rapple, 1994, p. 6). The Committee instructed examiners regarding how students were to be called for examination, including the manner in which students are to either stand or sit. Student knowledge of school subjects was to be examined in a specified order. The Committee also established rules for how students were to be questioned and how their answers were to be assessed (pp. 6-7).

### ***Expanding Role of Central Government in Education for Social Reproduction***

The prelude to the Revised Code of 1862 and its “payment-by-results” provision was characterized by an increasing financial and regulatory role of the British state into elementary education for the poor and working class. Rapple (1994) explains that previous, mostly religious, education efforts proved insufficient to provide schools for the growing number of poor children during the 19th century. Governing class concerns about the education of the lower classes spurred increased state intervention. The education of the poor and working class “was seen by some as an urgent humanitarian mission, and by others as a necessary undertaking if threats of revolution and moral decay were to be averted” (Coltham, 1972, p. 16). In the words of the Newcastle Commission of 1861, state efforts were geared toward “the education of the children of vagrants and to that of other children who cannot properly be allowed to associate with the families of respectable parents’.” (Rapple, 1994, p. 2; also see Coltham, 1972). Thus, increased state intervention into education corresponded with the official admission of a class-conscious theory that explicitly advocated a role for the state in the use of schooling for the reproduction of the social class order. As a result of these concerns, central government intervention increased as the 19th century progressed (Rapple, 1994; Coltham, 1972).

### ***The Demands of Empire: Control Expenditures on the Poor and Working Class***

As the 19th century progressed, the state increased expenditures on school buildings, the training of teachers, and the “capitation grant” or funds used to encourage attendance. Grants for books and equipment also appeared. “With all these expenses the amount of the grant voted each year necessarily grew,” Rapple (1994) explains, “until by 1859 it had risen to 723,115 pounds, not perhaps an inconsequential amount, but one which pales into some insignificance when set beside the nearly 78,000,000 pounds spent on the recent Crimean War” (p. 3). Nonetheless, the growth in expense on education meant, “the tentative period of state involvement was over” (p.3). And while the government was now committed to extending education, “it was a sine qua non that it be ‘cheap’ and especially so since the run on the coffers due to the Crimean War” (p. 3).

Lowe, the chief architect of payment-by-results, admitted that it was driven by concerns for what today is given as the demand to “balanced budgets.” Analyses of the Annual Reports of the Intermediate Education Board of Ireland during its nineteenth century “payment-by-results” system:

indicates that exam standards and rewards were systematically manipulated. As the percentage of students passing the exam increased, expenditures on performance rewards to teachers grew. When the pass rate became too high, and thus too costly, the tests were made more difficult and the standards for passing were increased in order to reduce pass rates. This manipulation of pass rates assured there was no significant upward or downward trend in the percentage of students passing during the results era. (Madaus, Russell Higgins, 2009, p. 121; see Madaus, Ryan, Kellighan, & Airasian, 1987)

In other words, the payment-by-results provision of the Revised Code was a means by which to contain educational costs. In England, the historical record indicates a substantial decline in average teacher salary during the initial period of payment-by-results with salaries only rising to previous levels by 1870; because this period also witnessed inflation, Rapple (1994) suggests that teachers were thus “affected severely”.

### ***Outcomes of “Payment-by-Results”***

The educational outcomes of this scheme were predicable, with critics even prior to its implementation foreshadowing present criticism of high-stakes testing and the so-called accountability movement. Historical accounts reveal that payment-by-results, wherever it was practiced, led to a scripted, low-level pedagogy; an over reliance on memorization, with a curriculum narrowed to examined subjects. Increased teacher turnover, exodus from the profession, and decline in the quality of teachers entering the profession during payment-by-results were additional outcomes noted in the historical record.

In most schools, in the early years of payment by results, teachers frequently received less from the annual grant than they previously earned as a fixed salary. As a consequence, many became peripatetic, changing their positions from school to school in search of greater remuneration. It was even said that, when seeking a position, some teachers calculated the percentage of passes in different school districts and were influenced by the scores in making their decision (Annual

Reports, 1871-72, p. 34). A multitude were sacked for securing poor grants; managers, declares Edmonds, “appointed and dismissed their teachers just as they ordered slates in preference to copy-books or vice-versa” (Edmonds, 1962, p. 77).

Many others, leaving teaching entirely, migrated to different occupations.... There is also much evidence that those who did enter teaching were of a lower caliber than formerly; their attainments, aims, and work habits frequently left much to be desired. (Rapple, 1994, p. 10)

Cheating, system gaming and abuse of students in an effort to secure their “performance” are additional documented outcomes of this system (Rapple, 1994).

After his retirement, one Inspector referred to the experiment as “an ingenious instrument for arresting the mental growth of the child, deadening all his higher faculties” (Coltham, 1972, p. 24). Coltham goes on to note that, “Critical educators of the time referred to ‘three-fourths of our schools’ as ‘grant factories’ and a child as ‘a grant-earning unit’ ” (p. 24). Far from “being for the kids” as present reformers would have it (e.g., <http://studentfirst.org>), this approach led government inspectors and teachers alike to oppose the system as anti-student, anti-teacher, anti-education and generally degrading to human dignity (Coltham, 1972; Dent, 1963; Griffiths, 1987; Madaus, Ryan, Kellighan & Airasian, 1987; Rapple, 1994).<sup>3</sup>

Yet, concern here rests not mainly with documenting the atrocities associated with the scheme, which are chronicled in detail by those cited above, but rather with the conditions that give rise to what appears, in light of this history, as irrational and obviously not the way to organize education.

### ***Origin of Performance Pay: Past and Present***

The origin of this scheme rests with efforts by elites to reduce or control expenditures on an expanding government administered education system. Importantly from the vantage of the present, the financial concern for economism was massive expenditures on the Crimean War, a war—like present wars continuously waged by the US against the popular will—considered to be a failure. That cost cutting is the driving force behind TBTEC is evident in the rhetoric of self-appointed reformers such as Bill Gates, along with academic allies such as Hanushek (e.g.,

2010). Gates (2011) argues that through TBTECT, we can “flip the curve,” that is decrease spending while raising performance. This line is repeated over and over again through numerous research briefs, foundations and think tanks that receive Gates (and other wealthy, neoliberal supporters) money. Even if test scores do not improve—as they have not in New York City (Fryer, 2011)—reformers appear to be intransigent: TBTEC will be imposed in one form or another, especially with the apparent support of a once critical union leadership (Dillon, 2011; Sawchuk, 2011).

Further, this accountability scheme, emerging in the context of governing class humanitarian concerns and fears of social revolution, served to increase supervisory authority of the national government over educational affairs, taking control away from religious authorities. Here too, a commonality with the present is found, as the expansion of central government power over education has been dramatically increasing, especially with NCLB and RTT. Just as central authorities overseen by Parliament gained control over the school curriculum by virtue of stipulating what curriculum would be tested, and thus what results would yield payment to educators, today’s reformers are crafting a national curriculum tied to national exams that will be used as the basis for state TBTEC policy, further limiting state autonomy.

Importantly, payment-by-results was born in a political context that did not emphasize or espouse education as the key means to opportunity or popular participation in government, as US mythology would have it. It emerged in the context of a society that held more to an open acceptance and reproduction of the class system, with the straightforward idea that schools should aim to prepare students for their station in life, something determined in large measure by their social origins. That is to say, payment-by-results originated in the context of a class-conscious system of education. It originated during a time where a national system of public education, with compulsory attendance, public oversight and public control and public finance did not yet fully exist. There is in the present situation evidence that the commitment to earlier notions of “opportunity” and preparation for civic life are being supplanted by more narrow renderings focusing on job training and bootstrapping for the poor, while private, elite offerings are provided for the rich (Glass, 2008; Kovacs, 2011).

### ***The Political Significance of TBTEC Under Obama***

Glass's (2008) contention is that leading reform proposals to improve education—charters, vouchers, and other school choice mechanisms—are in reality proposals to cheapen education for the poor and working class and privatize it for largely white, middle class. Yet, one of the few reforms to escape Glass's microscope is TBTEC. Nonetheless, his contention aligns with the historical record of payment-by-results in England and Ireland, and provides an important starting point for analyzing current TBTEC policies. It must be noted from the outset that tying teacher pay to student performance on examinations allows authorities to regulate and *quickly* adjust expenditures on teaching to meet any number of specific aims, without the hassle of previous arrangements which demanded a stability of funding, collective bargaining, due process, etc. That is to say, while unions may agree to build TBTEC into their contracts during collective bargaining (should collective bargaining not be outlawed) once evaluation, and ultimately pay, is determined by student test scores (irrespective of the technical merits of the model used), financial outlays for schools can be regulated by manipulating academic standards—as took place in the Irish case studied by Madaus, Ryan, Kellighan, and Airasian (1987)—*without teacher consent*.

### ***Basic Components of Current TBTEC Plans***

Throughout the 20th century, there have been sporadic experiments with TBTEC-like policies (Gratz, 2009; Johnson, 1984; Murnane & Cohen, 1986). Only in the past year, however, has there been a frantic move to adopt TBTEC on a national level. Inspired by the promise of increased federal funds through RTT if states adopted teacher evaluation policies favored by the US Department of Education, TBTEC systems are being quickly drafted in an emergency-like atmosphere in many states, and in some instances, these RTT spurred proposals contradict existing law. In Maryland, the state's RTT application contradicts a provision of the State's recently passed Education Reform Act, requiring that no single criteria—including student test scores—can contribute to more than 35 percent of a teacher's evaluation (Ujifusa, 2010).

While there are variations among the policies currently being adopted by states, common features are evident. In general, federal officials and the self-appointed billionaire neoliberal reformers advocate that teacher pay move from using *time* (seniority) and credentials as the criteria for salary to a system that evaluates and compensates teachers on various measures of “performance”. Thus there is a move away from fixed salary scales. The new approach calls for a

ranked evaluation system replacing the vilified binary system of “acceptable” or “unacceptable”. Despite the euphemisms and caveats (e.g., classroom observations) teacher quality, performance, etc., are defined as the degree to which a teacher’s students show improvement on state tests. (As a note, teacher educators will likewise be evaluated on the test scores the students of their students receive on state tests.) Finally, the ability to dismiss teachers on the basis of their students’ test scores—meaning the elimination of tenure and due-process related protections—is sought (this arrangement already exists in Washington, D.C., e.g., Lewin (2010) and Florida, with the *Student Success Act*, Sanders, 2011; see also Miner, 2010).

While much is made of the “bonuses” that teachers can receive as a result of increases in their students’ test scores, in the case of Florida, no funds have actually been allocated to pay for the bonuses (Grimm, 2011). In the case of Houston there was 16.5 percent decline between 2006 and 2010 in the number of teachers receiving a bonus; fewer teachers in schools serving large percentages of students living in poverty received a bonus as test scores strongly correlate with the social class of test takers. Teachers who have the most academically advanced students also received fewer bonuses as their students already score high on state tests, leaving little room for “growth” (Sarrío, 2011)<sup>4</sup>. These merit pay models eliminate salary schedules such that a teacher’s salary can only increase with test-based bonuses. This means it is possible for a teacher to show up every day for work, year after year, either struggling to help students living in poverty learn academic content, or working with students who already excel in school related subjects, and never receive even a cost of living increase to their base salary (if they are not fired outright).

### ***The Politics of “Value Added”***

Detailed analysis of the political significance of TBTEC can take place at two levels: the first examines the political content of TBTEC schemes themselves. This is the level of analysis presented here. The second level of analysis deals with the significance of the federal imposition of these schemes, which is the subject of another analysis of teacher evaluation as it relates to the Obama Administration reform agenda (Garrison, 2011b).

The method for using student test scores to evaluate teachers is called “value-added,” and has been intensely criticized by the scientific community (e.g., Baker et al, 2010). Yet, the political significance of these models has not been articulated. A prominent model, one now sold

to states and districts by the SAS Corporation

(<http://www.sas.com/govedu/edu/k12/evaas/index.html>), was first developed for Tennessee, and known as the Tennessee Value Added Assessment System (TVAAS) (Kupermintz, 2003).

One of the most basic criticisms leveled against these models is that improvements in student test scores cannot be proven to be an effect of a single school employee, nor can test score patterns be separated from the larger community of which students and educators are a part. Thus, Kupermintz (2003) articulates a key flaw in the logic of how TVAAS defines “teacher effectiveness,” and the ranking of teachers that ensues. These “value added” models imply, he writes, that there:

are two distinct variables—teacher effectiveness and differences in student learning—and that the former causes the latter. Unfortunately, such causal interpretation is faulty because teacher effectiveness is defined and measured by the magnitude of student gains. In other words, differences in student learning determines—by definition—teacher effectiveness: a teacher whose students achieve larger gains is the "effective teacher." TVAAS divides teachers into five "effectiveness" groups according to their ranking among their peers in terms of average student gains. To turn full circle and claim that teacher effectiveness is the cause of student score gains is at best a necessary, trivial truth similar to the observation that "all bachelors are unmarried." (p. 289)<sup>5</sup>

Understood politically, TVAAS is a technical justification for the power to define what effective teaching means. This vulgar operationalism aims to prove that teachers are responsible for patterns of educational outcomes that are in fact socially determined. Doing so institutionalizes an extreme individualism and “victim blaming” that is at the foundation of neoliberal ideology, and thus serves as a form of disinformation by decontextualizing “academic failure” from the larger society that produced the conditions for the failure in the first place. The elimination of seniority that is built into “value added” models is particularly significant. By eliminating seniority, the conditions are created for removing older, more experienced teachers. In addition to reducing expenditures on teacher salaries, this also removes a political barrier to implement the most draconian anti-educational practices, which younger teachers may not be as critical of

or in a position to combat. Further, parents who are seen by teachers as not contributing to students' test scores, will be positioned as standing in the way of their livelihood (this is predicted by the real history of payment-by-results). This challenges an already fragile relationship between teachers and parents, given the decades-long public campaign against teachers, which is now being exploited by “reformers” (e.g., Michelle Rhee’s group: <http://www.studentsfirst.org/>). The production of disunity between teachers and parents makes it easier for corporate reformers to maneuver and remain unaccountable.

“Value added” ranking of teachers is in essence norm-referenced: individual teachers are ranked against their school-system peers. But it is nonetheless unfair, as teachers working in the same state are nonetheless judged by different standards, as aggregate performance between districts varies greatly. Kupermintz explains: “Due to the substantial variability in performance between systems, teachers in low- and high-performing systems will be judged against very different evaluative criteria” (2003, p. 290). In addition to raising concerns of fairness, this mechanism creates incentives for teachers to compete against each other, and to view education as a competitive and individual, and not social, undertaking. Such practices serve to break unity, and thwart recognition of common interests and the development of collaborative efforts.

The results of tying rewards and sanctions to TVAAS are both nefarious and predictable, especially when viewed in light of the outcomes of payment-by-results. Kupermintz (2003) predicts that:

the confusion of student and teacher effects may unintentionally entice teachers to seek the strongest students at the expense of students who require more investment but can offer only small gains in return. As stakes increase, the incentive for teachers to target "high yield" students will intensify. Attaching tangible rewards or sanctions to value-added information is likely to encourage the development of a cynical calculus of the worth of different students to maximizing teachers' return on the investment.... The exclusive attribution of gains to teachers conceals potentially harmful practices whereby teachers are effective with certain students but not with others. Because student variability in gains within a classroom is averaged during the statistical calculations to produce a total teacher effect, teachers who tend to concentrate efforts on students who are

likely to demonstrate more robust gains at the expense of other, more challenging, students goes unnoticed. The negative long-term consequences of transforming student test score gains into the ultimate goal for teachers will probably be felt strongest by those students whom the new education legislation promised not to leave behind. Few teachers or schools would be able to afford ignoring the calculations of optimal gain in favor of pure pedagogical considerations. (p. 294)

Thus, embedded in TVAAS is a strong value preference for ranking human worth and the production of inequality. It also embodies an outlook that places value on risk taking with *other people's lives* for maximum *personal gain*. This anti-social value-set is clearly evident among corporate reformers who dream up, support and impose such schemes via their increasing control over political institutions.

### ***Conclusion***

This paper suggests that the impetus for TBTEC is the drive of a faction of the super rich to reduce the total financial expenditure on public schooling in the context of an expanding but faltering empire. The historical context for payment-by-results also reveals a political context not in favor of democratic purposes for education, and corresponds to a systemic disregard for the working class and poor. Does TBTEC in the present signal yet another move to a form of anti-democratic schooling? Those responsible for TBTEC, including President Obama and Arne Duncan, appear to envision a very narrow education devoid of even the pretense of democratic purpose. Their support for TBTEC cannot be supported by any reasonable interpretation of existing data, or the application of logic. Instead, TBTEC appears to be highly ideological and politically motivated, a means to limit the power of teachers (and administrators really) so as to remove a major block to the radical restructuring education in the US.

While this suggests many things, one thing is particularly importune for the critique of TBTEC: “value added” schemes cannot be effectively countered if only their technical failings are addressed. More accurately, the political content of TBTEC must be further articulated and exposed, so that these technical failings are not misunderstood as mere incompetence, ignorance or evidence of the limits of science as applied to education. These technical failings are in reality an outcome of the irrationality of the political project that stands behind TBTEC. To assist

educators and researchers in going beyond the technical limits of published critiques of TBTEC, this political project must be continuously exposed, and an alternative vision not limited to 19th century conceptions of public schooling should be articulated.

---

I would like to thank Agnieszka Zak-Moskal for her assistance with research for this article.

<sup>1</sup>The term reform is placed in scare quotes not simply to reveal my opposition to these efforts, but more importantly, to indicate that they are in fact not “reforms” but rather, as I have argued elsewhere (Garrison, 2009) deforms—that is, they are conscious efforts to wreck the existing public education system. Put simply, they should not be considered reforms because they have not nor will they serve to improve the quality of education because they take as their starting point a rejection of what is known about the theory and practice of education. Current reform efforts—whether under the banner of choice, accountability or closing the achievement gap—are best characterized as both anti-conscious (against thinking and careful analysis) and anti-social (that is, they reject outright the broadly social nature of education and its complex relationship to other social institutions, and reduce its purpose to the most narrow aim of job training and social control).

<sup>2</sup>Podgursky and Springer (2007) offer an American Enterprise Institute supported effort to render empirical support for performance pay that is nonetheless unconvincing: “While the literature is not sufficiently robust to prescribe how systems should be designed...it is sufficiently positive to suggest that further experiments and pilot programs by districts and states are very much in order” (p. 943). Of course, AEI and their fellow travelers, along with the Obama administration, have gone well beyond “pilot” programs that can be carefully assessed, as full scale TBTEC is built into federal policy and RTT competition and the administration’s *Blueprint for Education Reform*.

<sup>3</sup>These outcomes foreshadow in important ways abuses documented in the test-driven Knowledge Is Power Program (KIPP) and similar models of education praised by venture philanthropists and government officials (see Horn, 2009, 2011).

<sup>4</sup>Generally, the higher the previous achievement of students, the more likely their future teacher is to be rated effective by TVAAS, see Kupermintz (2003, p. 292).

<sup>5</sup>In fact, this is a long-standing theoretical error evident in much of the efforts to rank students on the basis of their intelligence—since intelligence is defined by success on tests of intelligence, a mere tautology is rendered as evidence of a cause and effect relationship—high scores on these tests are caused by intelligence! (see Garrison, 2009).

## References

- Baker, E. I., Barton, P. E., Darling-Hammon, L., Haertel, E., Ladd, H. F., Linn, R. L., et al. (August 29, 2010). Problems with the use of student test scores to evaluate teachers. Washington, DC: Economic Policy Institute.  
[http://epi.3cdn.net/724cd9a1eb91c40ff0\\_hwm6iij90.pdf](http://epi.3cdn.net/724cd9a1eb91c40ff0_hwm6iij90.pdf)
- Braun, B. (2010, May 10). Education Commissioner Schundler dismisses US test ranking N.J. at the top in reading, math. *Star-Ledger*.  
[http://blog.nj.com/njv\\_bob\\_braun/2010/05/us\\_education\\_tests\\_ranks\\_nj\\_at.html](http://blog.nj.com/njv_bob_braun/2010/05/us_education_tests_ranks_nj_at.html)
- Burton, E. F. (1979). Richard Lovell Edgeworth's education bill of 1799: A missing chapter in the history of Irish education. *The Irish Journal of Education, 13*(1), 24-33.
- Coltham, J. B. (1972). Educational accountability: An English experiment and its outcomes. *School Review, 81*(1), 15-34.
- Dent, H. C. (1963). The centenary of payment-by-results. *British Journal of Educational Studies, 11*(2), 185-188.
- Dillon, S. (2011, May 20). Former foes join forces for education reform, *New York Times*.  
[http://www.nytimes.com/2011/05/21/education/21rhee.html?\\_r=1&ref=samdillon](http://www.nytimes.com/2011/05/21/education/21rhee.html?_r=1&ref=samdillon)
- Emery, K., & Ohanian, S. (2004). Why is corporate America bashing our public schools? Portsmouth, NH: Heinemann.
- Finne, L. (2010). Bill Gates, Arne Duncan and Joel Klein criticize the way we pay teachers. Washington Policy Center. <http://www.washingtonpolicy.org/blog/post/bill-gates-arne-duncan-and-joel-klein-criticize-way-we-pay-teachers>
- Fryer, R. G. (2011). Teacher incentives and student achievement: Evidence from New York City Public Schools. Working Paper 16850. Cambridge, MA: National Bureau of Economic Research. <http://www.nber.org/w16850>
- Garrison, M. J. (2009). *A Measure of failure: The political origins of standardized testing*. Albany: SUNY Press.
- Garrison, M. J. (2011a). Clever rhetoric won't save your undemocratic reform from failure: An open letter to Arne Duncan on the occasion of teacher appreciation week.  
<http://www.markgarrison.net/archives/1084>
- Garrison, M. J. (2011b). The political economy of educational restructuring: On the origin of

- performance pay and Obama's "Blueprint" for education. In P. Carr & B. Porfilio (Eds.), *The phenomenon of Obama and the agenda for education: Can hope audaciously trump neoliberalism?* Charlotte, N.C.: Information Age Publishing.
- Glass, G. V. (2008). *Fertilizers, pills, and magnetic strips: The fate of public education in America*. Charlotte, N.C.: Information Age Publishers.
- Gratz, D. B. (2009). *The peril and promise of performance pay: Making education compensation work*. Lanham: Rowman & Littlefield Education.
- Griffiths, R. (1987). A tale of 804.573 horses: Arithmetic teaching in Victoria, 1860-1914. *Educational Studies in Mathematics, 18(2)*, 191-207.
- Grimm, F. (2011, March 19). Teacher merit pay is an illusion, *Miami Herald*.  
<http://www.miamiherald.com/2011/03/19/2122488/teacher-merit-pay-is-an-illusion.html>
- Hanushek, E. A. (2010), Cry wolf! This budget crunch is for real. *Education Week, 29*, 32.
- Hass, N. (2009, December 4). Scholarly investments. *The New York Times*. Retrieved from  
<http://www.nytimes.com/2009/12/06/fashion/06charter.html>
- Helper, S., Kleiner, M. M., & Wang, Y. (2010). Analyzing compensation methods in manufacturing: Piece-rates, time-rates, or gain-sharing? NBER Working Paper No. 16540. Cambridge, MA: National Bureau of Economic Research.  
<http://www.nber.org/papers/w16540>
- Horn, J. (2009). The KULT of KIPP: An essay review. *Education Review, 12(3)*.  
<http://edrev.asu.edu/essays/12n3index.html>
- Horn, J. (2011). Corporatism, KIPP, and cultural eugenics. In P. Kovacs (Ed.), *The Gates Foundation and the future of US "Public" schools* (pp. 80-103). New York: Routledge.
- Johnson, S. M. (1984). Merit pay for teachers: a poor prescription for reform. *Harvard Educational Review, 54(2)*, 175-183.
- Kovacs, P. (Ed.). (2011). *The Gates Foundation and the future of US "public" schools*. New York: Routledge.
- Kupermintz, H. (2003). Teacher effects and teacher effectiveness: A validity investigation of the Tennessee Value Added Assessment System. *Educational Evaluation and Policy Analysis, 25(3)*, 287-298.
- Lewin, T. (2010, July 23). School chief dismisses 241 teachers in Washington, *New York Times*.  
<http://www.nytimes.com/2010/07/24/education/24teachers.html>

- Madaus, G. F., Ryan, J. P., Kellighan, T., & Airasian, P. W. (1987). Payment-by-results: An analysis of a nineteenth century performance-contracting programme. *The Irish Journal of Education, 21(2)*, 80-91.
- Madaus, G. F., Russell, M. K., & Higgins, J. (2009). *The paradoxes of high stakes testing: How they affect students, their parents, teachers, principals, schools, and society*. Charlotte, N.C.: Information Age Pub.
- Martins, P. S. (2010). Individual teacher incentives, student achievement and grade inflation. <http://cee.lse.ac.uk/cee%20dps/ceedp112.pdf>
- Miner, B. (2010). Looking past the spin: Teach for America. *Rethinking Schools, 24(3)*. [http://www.rethinkingschools.org/restrict.asp?path=archive/24\\_03/24\\_03\\_TFA.shtml](http://www.rethinkingschools.org/restrict.asp?path=archive/24_03/24_03_TFA.shtml)
- Murnane, R. J., & Cohen, D. K. (1986). Merit pay and evaluation problems: why most merit pay plans fail and a few survive. *Harvard Educational Review, 56(1)*, 1-17.
- Peach, E. B., Wren, D. A. (1992). Pay for performance: From antiquity to the 1950s. In T. C. Mawhinney (Eds.) *Pay for performance: History, controversy, and evidence* (pp. 5-22), New York: Haworth Press.
- Perry, J. L., Engbers, T. A., & Jun, S. U. (2009). Back to the future: Performance-related pay, empirical research, and the perils of persistence. *Public Administration Review, January/February*, 39-51.
- Podgursky, M. J., & Springer, M. G. (2007). Teacher performance pay: A review. *Journal of Policy Analysis and Management, 26(4)*, 909-949.
- Rapple, B. A. (1994). Payment-by-results: An example of assessment in elementary education from nineteenth century Britain. *Education Policy Analysis Archives, 2(1)*, 1-21. <http://epaa.asu.edu/ojs/article/view/664>
- Ravitch, D. (2010). *The death and life of the great American school system: How testing and choice are undermining education*. New York: Basic Books.
- Sanders, T. (2011, March 24). Rick Scott signs teacher merit pay bill; critics calls it 'wrecking ball', *The Florida Times-Union*. <http://jacksonville.com/news/florida/2011-03-24/story/rick-scott-signs-teacher-merit-pay-bill-critic-calls-it-wrecking-ball>
- Saltman, K. J. (2010). *The gift of education: Public education and venture philanthropy*. New York: Palgrave Macmillan.

- Sarrio, J. (2011, April 3). In a Texas classroom, big lessons for Georgia. *The Atlanta Journal-Constitution*. <http://www.lexisnexis.com/hottopic/academic/>
- Sawchuk, S. (2011, May 13). NEA leaders propose teacher-evaluation shift, *Education Week*. <http://www.edweek.org/ew/articles/2011/05/11/31nea.h30.html?r=1119138732>
- US Department of Education. (2010). *A Blueprint for reform: The reauthorization of the Elementary and Secondary Education Act*. <http://www2.ed.gov/policy/elsec/leg/Blueprint/index.html>
- Ujifusa, A. (2010, November, 18). O'Malley eyes compromise on state teacher evaluations. *Maryland Community News Online*. [http://www.gazette.net/stories/11182010/prinsch175028\\_32542.php](http://www.gazette.net/stories/11182010/prinsch175028_32542.php)
- Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. The New Teacher Project. <http://widgeteffect.org/>
- Zhang, X. (2002). *Exploring the impact and process of merit pay for public school teachers*. Unpublished dissertation, The American University.