

**New York State Principals**  
**New York State's High Schools, Middle Schools and Elementary Schools**

*An Open Letter of Concern Regarding New York State's APPR Legislation  
for the Evaluation of Teachers and Principals*  
[www.newyorkprincipals.org](http://www.newyorkprincipals.org)

Over the past year, New York State has implemented dramatic changes to its schools. As building principals, we recognize that change is an essential component of school improvement. We continually examine best practices and pursue the most promising research-based school improvement strategies. We are very concerned, however, that at the state level change is being imposed in a rapid manner and without high-quality evidentiary support.

Our students, teachers and communities deserve better. They deserve thoughtful reforms that will improve teaching and learning for all students.

It is in this spirit that we write this letter, which sets forth our concerns and offers a path forward. We believe that it is our ethical obligation as principals to express our deep concerns about the recently implemented Annual Professional Performance Review (APPR) regulations. These regulations are seriously flawed, and our schools and students will bear the brunt of their poor design. Below we explain why we are opposed to APPR as it is presently structured.

***Background***

In May 2010, the New York State Legislature—in an effort to secure federal Race to the Top funds—approved an amendment to Educational Law 3012-c regarding the Annual Professional Performance Review (APPR) of teachers and principals. The new law states that beginning September 2011, all teachers and principals will receive a number from 0-100 to rate their performance. Part of that number (ranging from 20% to 40%) will be derived from how well students perform on standardized tests. At first glance, using test scores might seem like a reasonable approach to accountability. As designed, however, these regulations carry unintended negative consequences for our schools and students that simply cannot be ignored. Below we explain both the flaws and the consequences.

***Concern #1: Educational research and researchers strongly caution against teacher evaluation approaches like New York State's APPR Legislation***

A few days before the Regents approved the APPR regulations, ten prominent researchers of assessment, teaching and learning wrote an open letter that included some of the following concerns about using student test scores to evaluate educators<sup>1</sup>:

- a) *Value-added models (VAM) of teacher effectiveness do not produce stable ratings of teachers. For example, different statistical models (all based on reasonable assumptions) yield different effectiveness scores.<sup>2</sup> Researchers have found that how a teacher is rated changes from class to class, from year to year, and even from test to test<sup>3</sup>.*
- b) *There is no evidence that evaluation systems that incorporate student test scores produce gains in student achievement. In order to determine if there is a relationship, researchers recommend*

---

<sup>1</sup> Baker, E. et al. (2011). Correspondence to the New York State Board of Regents. Retrieved October 16, 2011 from: [http://www.washingtonpost.com/blogs/answer-sheet/post/the-letter-from-assessment-experts-the-ny-regents-ignored/2011/05/21/AFJHIA9G\\_blog.html](http://www.washingtonpost.com/blogs/answer-sheet/post/the-letter-from-assessment-experts-the-ny-regents-ignored/2011/05/21/AFJHIA9G_blog.html).

<sup>2</sup> Papay, J. (2011). Different tests, different answers: The stability of teacher value-added estimates across outcome measures. *American Educational Research Journal* 48 (1) pp 163-193.

<sup>3</sup> McCaffrey, D. et al. (2004). Evaluating value-added models of teacher accountability. Santa Monica, CA.; Rand Corporation.

*small-scale pilot testing of such systems. Student test scores have not been found to be a strong predictor of the quality of teaching as measured by other instruments or approaches*<sup>4</sup>.

- c) *The Regents examinations and Grades 3-8 Assessments are designed to evaluate student learning, not teacher effectiveness, nor student learning growth*<sup>5</sup>. Using them to measure the latter is akin to using a meter stick to weigh a person: you might be able to develop a formula that links height and weight, but there will be plenty of error in your calculations.

**Concern #2: Students will be adversely affected by New York State's APPR**

When a teacher's livelihood is directly impacted by his or her students' scores on an end-of-year examination, test scores take front and center. The nurturing relationship between teacher and student changes for the worse.

- a) *With a focus on the end of year testing, there inevitably will be a narrowing of the curriculum as teachers focus more on test preparation and skill and drill teaching*<sup>6</sup>. Enrichment activities in the arts, music, civics and other non-tested areas will diminish.
- b) *Schools will have an incentive to place struggling students in lower-level classes without standardized assessments. School systems may hesitate placing students in Regents classes beyond the basic five needed for graduation so that their performance on Advanced Regents examinations will not negatively impact evaluations. If schools use Advanced Placement (AP) or International Baccalaureate (IB) scores, as suggested by Commissioner King, schools might be more reluctant to challenge students upward for fear that poor test performance might result in teachers being unfairly penalized.*
- c) *Teachers will subtly but surely be incentivized to avoid students with health issues, students with disabilities, English Language Learners or students suffering from emotional issues. Research has shown that no model yet developed can adequately account for all of these ongoing factors*<sup>7</sup>.
- d) *The dynamic between students and teacher will change. Instead of "teacher and student versus the exam," it will be "teacher versus students' performance" on the exam.*
- e) *Collaboration among teachers will be replaced by competition. With a "value added" system, a 5th grade teacher has little incentive to make sure that her incoming students score well on the 4th grade exams, for incoming students with high scores would make her job more challenging. When competition replaces collaboration, every student loses.*

---

<sup>4</sup> See Burris, C. & Welner, K. (2011). Conversations with Arne Duncan: Offering advice on educator evaluations. *Phi Delta Kappan* 93 (2) pp 38-41.

<sup>5</sup> New York State Education Department (2011). *Guide to the 2011 Grades 3-8 Testing Program in English Language Arts and Mathematics*. Retrieved October 18, 2011 from <http://www.p12.nysed.gov/apda/ei/ela-math-guide-11.pdf>.

<sup>6</sup> Committee on Incentives and Test-Based Accountability in Education of the National Research Council. (2011). *Incentives and Test-Based Accountability in Education*. Washington, D.C.: National Academies Press.

<sup>7</sup> Baker, E. et al (2010). Problems with the use of test scores to evaluate teachers. Washington, D.C. Economic Policy Institute. Retrieved October 16, 2011 from: [http://epi.3cdn.net/b9667271ee6c154195\\_t9m6iij8k.pdf](http://epi.3cdn.net/b9667271ee6c154195_t9m6iij8k.pdf);

Newton, X. et al. (2010). Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts. *Education Policy and Analysis Archives*. Retrieved October 16, 2011 from <http://epaa.asu.edu/ojs/article/view/810/858>.

; Rothstein, J. (2009). Student sorting and bias in value-added estimation: Selection on observables and unobservables. *Education Finance and Policy*, 4(4), 537-571.

### ***Concern #3: Tax Dollars Are Being Redirected from Schools to Testing Companies, Trainers and Outside Vendors***

School budgets must now be built within the constraint of New York State's 2% Tax Cap law. Due to the details of the APPR regulations, district funds must be funneled to staff development and outside scoring even as New York State taxpayers' precious dollars are funneled to testing companies and other vendors. At a time of economic crisis, this leaves fewer and fewer dollars for our classrooms. According to a recent report by the New York State Superintendents entitled, *At the Edge*<sup>8</sup>, 81% of all superintendents are worried that they will not have the funds to implement APPR in a way that would best serve their students' needs.

- a) *As recommended by the Regents reform agenda, all building principals receive ten days of training and superintendents receive seventeen days of training during the first year of implementation. Taxpayer funds to these outside training sessions are typically \$120 per day per person. Conservative estimates put the state-wide mandated costs for this training in the range of several millions of taxpayer dollars. Recognizing that many other supervisors will be required to attend this training in order to evaluate teachers, these costs will continue to balloon.*
- b) *The need for a rapid turnaround in test scores (so they can be included in end-of-year evaluations) coupled with an increased emphasis on exam security will result in increased costs for districts. Tax dollars will be diverted to outside companies in the areas of test development, exam security and data analysis. These dollars diverted to testing companies may well range into the hundreds of millions of dollars statewide at the same time that funding for instruction is cut.*

### ***Our Recommendations***

#### **1. School-wide achievement results should be used as part of every teacher's and principal's evaluation.**

Schools are learning communities. An evaluation system that even partially bases an individual teacher's evaluation on his or her students' scores ignores the reality that student success is often predicated on the work of many adults in a school, as well as out-of-school factors that are actually more responsible for student success<sup>9</sup>. Non-classroom-teacher factors account for roughly 85-90% of the statistical variation in students' test scores<sup>10</sup>. Reading teachers, resource room teachers and ELL teachers provide critical skill support to our students. Guidance counselors, social workers, psychologists and deans provide social-emotional support for students. Thus, student scores reflect the work of many school personnel, as well as the influence of families and the effort of students. Schools operate best when there is cooperation among all faculty members and when all are accountable for every student's learning<sup>11</sup>. If students' test scores must be used as part of educators' evaluations, we recommend that a score be used in a global manner. That is, that one index which incorporates school-wide achievement results be developed and used as part of each teacher's and principal's evaluation. Districts, based on the needs of their students, should have the final say in what achievement factors comprise that index.

---

<sup>8</sup> The Council of School Superintendents. (2011). *At the edge: A survey of New York State School Superintendents on fiscal matters*. Retrieved on October 16, 2011 from: <http://www.nyscoss.org/pdf/upload/AttheEdgeSurveyReportFINAL.pdf>

<sup>9</sup> Goldhaber, D., Brewer, D. & Anderson, D. (1999). A three-way error components analysis of educational productivity. *Education Economics* 7 (3) pp 199-208.

<sup>10</sup> See Hanushek, E., Kain, J. & Rivkin, S. (1998). *Teachers, schools and academic achievement*. Retrieved October 16, 2011 from [http://www.cgp.upenn.edu/pdf/Hanushek\\_NBER.PDF](http://www.cgp.upenn.edu/pdf/Hanushek_NBER.PDF).

<sup>11</sup> DuFour, R., & Eaker, R. (1998). *Professional learning communities at work. Best Practices for enhancing student achievement*. Bloomington, IN: National Educational Service.

**2. Pilot and adjust the evaluation system before implementing it on a large scale.**

Any annual evaluation system should be piloted and adjusted as necessary based on field feedback before being put in place state-wide. In other words, the state should pilot models and then use measures of student learning to evaluate the model. Delaware spent years piloting and fine-tuning their system before putting it in place formally state-wide.

**3. Use broad bands, not numbers, for the evaluation of teachers and principals.**

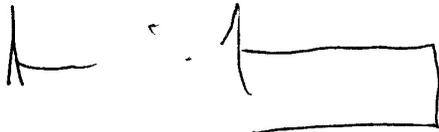
Evaluation should be about improvement, not sorting and selecting. A number between one and one hundred simply cannot describe the complex work of an educator. Neither is it realistic to think that the difference between a teacher score of 86 and 87 would have meaning or validity. The four performance bands (ineffective, developing, effective and highly effective) are the only evaluation ratings that should be used.

We, Principals of New York State schools, conclude that the proposed APPR process is an unproven system that is wasteful of increasingly limited resources. More importantly, it will prove to be deeply demoralizing to educators and harmful to the children in our care. Our students are more than the sum of their test scores, and an overemphasis on test scores will *not* result in better learning. According to a nine-year study by the National Research Council<sup>12</sup>, the past decade's emphasis on testing has yielded little learning progress, especially considering the cost to our taxpayers.

We welcome accountability and continually strive to meet high standards. We want what is best for our students. We believe, however, that an unproven, expensive and potentially harmful evaluation system is not the path to lasting school improvement. We must not lose sight of what matters the most—the academic, social and emotional growth of our students.

***Prepared by***

Sean C. Feeney, Ph.D.



President  
Nassau County High School Principals' Association  
Principal, The Wheatley School

Carol C. Burris, Ed.D.



Principal  
South Side High School

---

<sup>12</sup> Committee on Incentives and Test-Based Accountability in Education of the National Research Council. (2011). *Incentives and Test-Based Accountability in Education*. Washington, D.C.: National Academies Press.