



Testimony to the  
Senate Standing Committee on Education  
The Regents Reform Agenda:  
“Assessing” Our Progress

Mark D. Beehler  
Vice President  
Empire State School Administrators Association

October 16, 2013  
Buffalo, New York

Senator Flanagan and honorable members of the committee, my name is Mark Beehler. I am the Chief Information Officer and Director of Science for the West Seneca School Central District. I also serve as the President of the Western New York Administrators Association and as a Vice President of the Empire State School Administrators Association.

On behalf of the Empire State School Administrators Association and the 3,400 school administrators we represent, thank you for investing the time and resources into these opportunities to hear testimony from educators and school related personnel on the challenges and successes of the Regents Reform Agenda. It is apparent that concerns from the field are being heard. The recent release of approximately 25% of the ELA & Math 3-8 assessment questions and the loosening of the pre assessment requirement for Student Learning Objective's is evidence of NYSED's willingness to respond to the needs of educators and students. We hope and anticipate more positive changes will result from these hearings.

The past three years have born witness to unprecedented numbers of reform initiatives, regulation changes, curriculum shifts, and assessment requirements. Virtually every facet of education has been modified or completely overhauled. During this time, funding for schools has been reduced significantly and the ability to compensate and raise funds locally has been compromised. The loss of basic/foundation aid drastically reduces the ability of local education agencies to respond to community needs in a thoughtful, meaningful, and successful manner. The result is misunderstanding and resistance.

We are not opposed to being held accountable to high standards and we strive for continuous improvement in our practice however, the sheer volume of these changes has resulted in a cluster of 'initiatives' which are being implemented simultaneously without a deep understanding of their affect on one another. While some of these initiatives have the potential to help students and teachers, others do not. It is critical that we take time now to disambiguate this cluster of initiatives, identify the true nature of each initiative, articulate the manner in which each should be implemented, and modify those that do not work or are in conflict with the goals of public education in a democracy.

My testimony will focus on two primary areas of concern.

- The linking of student test performance with teacher and principal evaluation
- The loss of local control over curriculum, instruction, and information

## **Student Test Performance and Teacher/Principal Evaluation**

The linking of student test performance scores to teacher effectiveness appears to make sense on the surface. If we believe what research tells us- the influence of teachers is the single-most important (in school) factor in determining student achievement (Marder, 2012), then it would make sense to use that achievement data to evaluate teachers. New York's education law 3012-c requires school districts to include student performance for 40% of a teacher's performance rating. Depending upon your teaching assignment or the type of school building you are the principal of this may be comprised of a state growth score or a local Student Learning Objective. Inherent within this model is disparity between teachers and principals who receive a state assigned growth score and those that do not. This finding is consistent with the conclusions drawn from a study of the Tennessee Teacher Evaluation System, a system which appears to be a model for the New York system. In this study, teachers who had the choice of selecting the measures and targets for their students, those who would be writing Student Learning Objectives in New York, chose measures that would yield the highest growth score, and not necessarily measure appropriate student learning. This disparity in growth scores, state assigned or locally developed, is now a reality in most schools in New York and a serious flaw in the evaluation system. This widely recognized inequity is seen as unfair and is lowering the quality of performance and frustrating those subject to it.

Tying student test performance to teacher and principal evaluation systems creates a high stakes testing environment where the consequences of student success not only affect the student but the teacher and principal as well. This increased pressure only exacerbates an already stressful system precipitating all the consequences that research tells us will occur: a narrowed curriculum; teaching to the test- teaching discrete facts that are isolated from their context and focusing on rote memorization; and a decrease in the validity of the assessment tool. Even when tests are designed to assess higher order thinking skills, critical thinking abilities, and/or incorporate complexity of text and tasks, these negative outcomes will occur, and the tests themselves loss validity.

World renowned social psychologist Donald Campbell, a leader in quantitative experimental research concluded that the more any quantitative social indicator (or even some qualitative 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). Given the toxic environment that assessments are administered in, the validity and reliability of the assessments are rendered suspect. What is now known as Campbell's Law explains the increase in testing scandals across the country causing states to invest even more funding and excessive regulations to counteract the affects of a policy that should be abandoned for assessment systems that can help inform and improve education.

Additionally, New York has implemented the curriculum-testing system backwards. Testing is done to measure the depth and breadth of learning of a specific body of knowledge and/or set of skills. You first teach the knowledge and skills with a well developed curriculum and then sample from it with an assessment to measure progress. The recent administration of common core aligned assessments was done knowing that students were not prepared with a well developed curriculum. We knew that many more students would fail. Why would we deliberately create an environment where students would fail or at best be less successful than possible? Not only were students harmed by this process but it damaged the evaluation process fostering mistrust and necessitating adjusted performance levels and scaling by percentiles to compensate for APPR ramifications of the massive failure rates.

Schools are now faced with a game of catch up. Nearly every school across the state, particularly urban schools, have significant increases in the number of students who require Academic Intervention Services (AIS) for not meeting proficiency with respect to material they were not taught. Most schools are now diverting resources to remediate students instead of focusing those resources on delivering high quality initial instruction. The number of students who are mandated to receive these services has been decreased by adjusting the cut points for requiring services. But while this assists schools who simply do not have the staffing to meet the needs of all the students, these are still students who are not proficient. If we believe our AIS assists students who are not proficient and know the tests will be just as rigorous this year aren't we setting up these students for another year of failure?

I have heard the analogy of building the plane while it is flying. I understand its meaning. I also understand why, as a society, we have chosen not to do it. It is not a good idea. Planes are designed by engineers to meet a specific purpose, they follow specific regulations ensuring the safety and consistency of production, budget limitations are addressed, the effects on the environment are considered. Then safety considerations are made- a lot of them. Finally, it is tested in phases by a limited number of people. We do not simply load up the first aircraft off the assembly line with thousands of people and hope for the best. Yet this is how many of our teachers and administrators are feeling. Additionally, while planes are built on the basis of the knowledge, skill and wisdom of engineers, current practice excludes the very people who are to do the heavy lifting from taking part in crafting the curriculum and assessments that are the mainstay of education work.

There is a positive component of the testing system that has arisen from APPR but it has been tainted by its connection with teacher and principal evaluation, the new metric of student growth. I believe this metric has altered the role of the classroom teacher and provides a high level of insight into student learning. Under No Child Left Behind legislation and likely since the beginning of testing, teachers and students determined success by meeting a proficiency point, by passing the test. Demonstrating growth is now just as important as passing the test. This new metric of growth measures the learning as opposed to simply identifying the achievement rate of a student. Unfortunately, the link between the growth score of a student and the growth score

assigned to teachers and principals causes mistrust which is compounded by the failure of the state to release adjusted student growth scores for students.

Growth scores can be very insightful for educators but it is important to recognize that growth scores are a ranking. They are not a measurement. Growth scores are being misused in our APPR system and are purported to measure teacher effectiveness. They do not. They are simply a basis for comparing students to one another and subsequently, for comparing teachers and principals to one another. As useful as this may be, it measures nothing. For example, if five students were asked to arrange themselves in order of increasing height even if it is done correctly we still have no idea how tall any of them are. We only know how tall one is compared to the other. This is much different than actually measuring the height of the student with a measurement tool.

I indicated that research tells us the influence of teachers is the single-most important (**in school**) factor in determining student achievement. Those two words in the parenthesis are critically important because as this statement is being prepared, the APPR evaluation data has been released for Syracuse and Rochester teachers and principals- 40% of teachers are not effective. Unfortunately, the single most consistent predictor of student achievement, the student's family income (Hursch, 2005), has migrated through the growth scores and now affects the APPR rating of teachers and principals. What is most alarming is that we already knew this would happen! Page 1 of the American Institutes for Research growth model technical report is states: "Despite the model conditioning on prior year test scores, schools and teachers with students who had higher prior year test scores, on average, had higher MGPs. **Teachers of classes with higher percentages of economically disadvantaged students had lower MGPs**" (American Institutes for Research, 2012, p. 1).

### **The loss of control over curriculum, pedagogy, and information**

The Board of Regents adopted the Common Core Learning Standards in 2010 but its implementation was challenged greatly by reduced funding and the overlapping introduction of APPR requirements which took precedent over the learning standards. The ability of a school district to respond to such a major curriculum shift by identifying underlying standards, aligning existing resources, developing and/or purchasing new instructional resources, and providing the critical staff development and training to multiple disciplines simultaneously was more than most school districts had the capacity to handle. This was particularly true when districts also needed to retrain nearly the entire staff on new evaluation rubrics.

The recent administration of Common Core aligned Math and ELA assessments and their poor results is evidence that the curriculum was not fully implemented and successfully executed in a skillful manner. This is not because our teacher and principals have chosen to ignore them but

simply because too much change occurred in this small timeframe. We recognize the need to make changes but are simply not capable of adapting to so much change at one time.

With the limited resources, compressed timeframe, and high stakes associated with Common Core implementation many districts have been placed in a position where adopting the New York State Modules for ELA and Math is the only option. These modules are scripted plans that automate the teaching process and drastically limit the ability of local school districts to reflect the communities they serve by adjusting curriculum. Additionally, these modules prescribe a specific pedagogy not only instructing teachers exactly what to say and when but exactly how to say it. I fear that when parents recognize how drastically their schools resemble automated factories focused on turning out students as products they will not be happy. It is apparent that value has been placed on specific learning content, instructional pedagogy, and assessment style that favors particular students and excludes the innate talent and strengths of many other students thereby creating an uneven playing field for all of our students. This value judgment has not occurred at the local education agency level, and thus it serves to remove control from parents, school boards and elected official oversight.

As the Chief Information Officer for the district, I coordinate the collection, aggregation, stewardship, and dissemination of hundreds of thousands of data points for students and teachers. My role in the school system has changed drastically over the past few years and I have had the opportunity to see New York's transition from collecting two small database files of student information ten years ago to the coding of students with unique identifiers and real time data harvesting. What is greatly disappointing are the incredible resources that are expended in the collection, sorting, verifying, and distribution of data. This data is exceptionally important and CIO's across the state take the stewardship of data very seriously.

The recent publicity surrounding the collection of student information by InBloom is a concern for myself and many of my colleagues. We know that FERPA is the baseline for the protection of student information and many school districts have developed even more stringent guidelines over the dissemination of student information. What is troubling is that InBloom is aggregating student information on a massive scale without the consent of school districts. School superintendents have never been consulted or asked to provide consent for the collection of this data. More importantly, parents have never been asked for consent nor have they been formally informed of the disclosure practices for their own children's information.

It is important to understand the flow of data through the New York system. Student data originates with local education agencies and their student management system. Periodically data is extracted from the student management system and placed into a computer container at their Regional Information Center (RIC). This computer container is called Level 0 and is designated for the scrubbing and matching of student information. Information from various systems is loaded into Level 0 and matched up. CIO's correct errors in dates, track down missing students, insert missing data elements, etc. When the data is clean it is sent to another computer called

Level 1 where student assessment information is matched up. Level 1 is also called the Data Warehouse where educators can run various reports to help inform instruction. Throughout the year, there are reporting requirements and information is pulled from Level 1 and migrated to Level 2- for New York State Reporting. InBloom is pulling information from Level 2.

As a requirement for receiving RTTT funding, school districts must choose a data dashboard or portal from one of the three approved vendors. The data dashboard will take the information from InBloom and present it to parents and educators in a useful and meaningful manner. This all sounds great until the RTTT funding for the dashboard runs out and school districts must now bear the burden of the cost.

But the majority of a school district's cost for the use of the data dashboard does not come from the vendor who actually creates the data dashboard. It comes from the fees charged by InBloom to give the vendor the data. The data that it received from Level 2, which came from Level 1 after being cleaned and verified by school district employees in Level 0 which was collected by school district employees in their student management system. So schools must now buy back the data they provided at their own expense from InBloom at a cost of between \$2.00 and \$5.00 per pupil per year.

I recently spoke with one of the data dashboard vendors and learned they can give us the same product with Level 1 data meaning they do not need InBloom to produce a data dashboard product. Why must RTTT funding be expended on data that is already present in our RICs? Why must every district in the state that cooperated and agreed to RTTT pay InBloom for a service that could be provided by the existing data flow structure? Why must between \$5.2 million and \$13.5 million dollars be paid to InBloom every year for something that can be done already by our existing structure? for data that we already have? Couldn't we reinvest those funds into our existing structure to build systems that more accurately meet local needs?

## **Summary**

We are now several years into the Regents Reform Agenda. This hearing to assess our process affords an opportunity to identify not only how far we have come, but also how much farther we need to go, identify what we need to change, and ask 'Do we really need to go in that direction or this fast?' As reform initiatives are imposed that seek to improve the quality of education and ensure our students are college and career ready, we must value our schools for the complex systems they are and respect the communities they reflect. We are now at point where it is critical to regroup, renew or revise our mission, and ensure we have appropriate resources to move forward; or risk reducing or losing many of the advances that have occurred over the past several years.

On behalf of the Empire State School Administrators Association, I thank you for this opportunity to share.

## Reference List

American Institutes for Research (2012). 2011-2012 Growth model for educator evaluation technical report: FINAL. Retrieved from: <http://www.engageny.org/resource/2011-12-technical-report-for-growth-measures>

Campbell, D. T. (1976). Assessing the impact of planned social change, *The Public Affairs Center, Dartmouth College, Hanover New Hampshire, USA*.

Hirsch, D. (2005). The growth of high stakes testing in the USA: accountability, markets, and the decline in educational equality, *The British Educational Research Journal*, 605-622.

Marder, M. (2012). Measuring teacher quality with value-added modeling. *Kappa Delta Pi Record*, 48, 156-161.