

The following are the duly deliberated and enacted positions of the Board of Education of the New Paltz Central School District.

**On the Value Added Model (VAM):**

1. There are no studies that validate the efficacy of NYSED's Value Added Model, and numerous scientific studies have invalidated the efficacy of Value Added Models. Therefore, we oppose the expenditure of taxpayer funds and use of class time on tests and scoring systems where empirical evidence does not support their use.<sup>1</sup> We also oppose the use of New York's student population as test subjects, year after year, in the hope that doing so will eventually generate such evidence. Our primary obligation is the education of our children -- they are not test subjects for future generations.
2. There is a belief within the state education administration and among elected state officials that the evaluation model could produce a bell curve of teacher effectiveness.<sup>2</sup> It is our position that the sample group -- professionally certified educators subject to continuing professional development -- would not exhibit a normal distribution, negating the bell curve expectation. The focus on tinkering with pre-assessment metrics, sometimes called the "black box" (the base of expectations regarding student growth) is a deliberate manipulation of inputs in order to produce preordained outcomes. Even if the VAM were known to be efficacious its results would have to be considered in real terms, not in terms of how they failed to meet expectations. The fact that the use of the VAM in New York State is driven by expectations, rather than outcomes, demonstrates that it must not be used. Only objective, valid, and reliable assessment measures are acceptable.
3. Two very important criteria in measurement quality are reliability and validity. Reliability means the quality of the measurement methods suggest that the same data would have been collected each time in repeated observations of the same phenomenon. Validity means the measurement accurately reflects the concept it is intended to measure. One does not ensure the other, and neither are present in the VAM methodology. Recent results reveal that state growth scores fluctuate wildly from year to year among the same teachers and within the same districts.<sup>3</sup> This would not happen if the assessment model were working as a valid and reliable measure of teacher effectiveness. Results to date are too volatile and random considering the relatively constant demographics and short time frame. The lack of an explanation for the volatility indicates that the measures fail the reliability and validity criteria.
4. The practice of changing the "cut scores"<sup>4,5</sup> when the tests do not produce sufficient failures -- or produce an abundance of successes -- demonstrates that the tests are not objective measures. Rather, they are being manipulated to fulfill pre-assessment projections, not post-assessment analysis intended to inform improvements in education.
5. The pre-assessment metrics are not available for review by independent experts and school districts. The secretive nature of their contents and weights mean that districts cannot use the data to inform the improvement and delivery of education and the year-to-year variability precludes any longitudinal use.
6. State growth scores are economically biased. In districts with 90% or more economically disadvantaged students, 19.1% of educators are rated on state growth scores as developing or ineffective. This compares to 5% of teachers in schools with less than 40% economically

disadvantaged students. This indicates that environment is a significant, independent driver of state growth scores and a clear disincentive for educators to work in the districts most in need of great teachers.<sup>6</sup>

7. Test refusals have added to the volatility of the system. A simulation of statewide student test scores places 13% of educators' average student scores at the "extremes" (below 1.9 or above 2.4) when all students take the test. When refusals grow to 20%, as they sit at the moment, those extreme cases grow by 34%. When 50% of students refuse -- a reality for 122 districts statewide -- those extreme cases grow by 103%. Parental action taken to protect the welfare of their children should not adversely affect our educators.<sup>7</sup>

Our conclusion is that the results produced by the current assessment system are unproven, volatile, and lack utility. We call upon the Board of Regents and Legislature to immediately suspend all state assessments that use a VAM or growth theory until there is evidence of efficacy.

### **On the Current APPR mandates:**

1. For the reasons enumerated above it is inappropriate to use results of the current state assessment regimen for any APPR purpose.

2. Changes in the impact of test results on the tenure process -- and on teacher removal -- are mandated, which creates legal ramifications that are not connected to district-wide teacher quality improvements. An ineffective educator who has been fortunate to get acceptable test results due to the randomness inherent in the system cannot be removed for instructional ineffectiveness. Probationary educators will be prohibited from receiving tenure, even though the test results during their first years of employment provide no information on how they might improve. These changes conspire to mis-classify educators and rob school districts of local control.

3. Since 2010, when the new systems were put in place, the number of education majors in New York State has fallen by 40%.<sup>8,9</sup> States throughout America are already experiencing new teacher shortages. Reductions in job security and capricious nature of career-determining assessments have contributed substantially to the decline. The future availability of a sufficient number of certified teachers is in doubt.

4. The use of state assessment tests to evaluate teaching and learning corrupts the nature of instruction,<sup>10,11,12</sup> and has never been shown to increase student achievement. Teachers are no longer able to meet the needs of individual students as state assessments dictate the pace of instruction for the entire class while ignoring economic, geographic, and cultural differences.

Our conclusion is that the current APPR mandates are invalid measures of educator- and school district-effectiveness and present serious short- and long-term risks to the availability of instructional talent.

### **On the Utility of State Assessment Data:**

1. Student tests, when appropriately designed, are valid measures of student achievement,

but they are not efficacious measures of teacher effectiveness.

2. The 3-8 state assessments and Regents Aspirational Performance Measure (APM) are both intended as assessments of college and career readiness. However, 29.6% of our 8th grade students statewide demonstrate proficiency on the state assessments<sup>13</sup> vs. 37.5% grading as proficient via the Regents APM.<sup>14</sup> The 26% increase within such a short time is clear evidence of a lack of longitudinal alignment and reinforces our position that the 3-8 state assessments do not measure what is claimed.

3. The assessments do nothing to address the most pressing issue affecting academic performance: poverty. A non-economically disadvantaged student whose district is in the 10th percentile of economic disadvantage (affluent) is 310% more likely to achieve proficiency than an economically disadvantaged student whose district is in the 90th percentile.<sup>15</sup> State assessments do nothing to address that issue other than exacerbating the problem by diverting time and money away from real solutions.

Our conclusion is that the data produced by the state assessment system provide no value while simultaneously diverting resources away from initiatives that serve districts' missions.

### **Requested Actions:**

The Board of Education of the New Paltz Central School District asks the Board of Regents, State Education Department, New York State Legislature, and Governor Andrew Cuomo to declare an immediate moratorium on the current testing mandates and for that moratorium to continue until such time as a body of evidence for their efficacy in improving instruction has been fully established. We also request that no Smart Bond funds are expended to computerize an evaluation system based on the Value Added Model. Along with the aforementioned governing bodies we are copying this request to our local Regent Josephine Finn, Commissioner Elia, Assembly Speaker Heastie, Senate Majority Leader Flanagan, our State Senators John Bonacic and George Amedore, and Assemblyman Kevin Cahill with the expectation that they will take appropriate action in the interest of our children, the teaching profession, the taxpayers, and the future of public education in New York State.

### **Citations and Notes:**

1. ASA Statement on Using Value-Added Models for Educational Assessment. (2014).  
[http://www.amstat.org/policy/pdfs/asa\\_vam\\_statement.pdf](http://www.amstat.org/policy/pdfs/asa_vam_statement.pdf)

2. Woodruff, C. (2015, August 31). L.I. teacher's lawsuit on evaluation rating is microcosm of issues of APPR fairness. On Board Online.  
<http://www.nyssba.org/news/2015/08/27/onboardonlineaugust312015/l.i.teacherslawsuitonevaluationratingismicrocosmofissuesofapprfairness/>

3. Based on a comparison of intra-district growth scores present in the 2013-14 and 2012-13 Teacher Evaluation Databases. Data source: <http://data.nysed.gov/downloads.php>

4. Memo: "Interpreting Scores on the 2013-14 New York State Alternate Assessment". NYSED, October, 2014.  
<http://www.p12.nysed.gov/assessment/nysaa/201415/nysaainterpretingscores14.pdf>

5. An analysis of raw score to scale score conversion charts reveals changes over time. For example, the percentage raw score necessary for proficiency on Grade 5 ELA has increased from 2013-2015 as follows: 68%, 70%, 74%. Grade 4 Math changes: 65%, 71%, 68%. "English Language Arts (ELA) and Mathematic Assessment Results". NYSED. <http://www.p12.nysed.gov/irs/elamath/>
6. Based on an analysis of growth scores present in the 2013-14 and 2012-13 Teacher Evaluation Databases. Data source: <http://data.nysed.gov/downloads.php>
7. Proprietary simulation utilizing the following criteria: 25 students per teacher, 12 scores necessary to issue a score for a teacher, score distribution identical to statewide results. 100,000 teachers (2.5 million students) simulated per refusal level.
8. Sawchuk, S. (2014, October 22). Steep Drops Seen in Teacher-Prep Enrollment Numbers. Education Week. <http://www.edweek.org/ew/articles/2014/10/22/09enroll.h34.html>
9. Title II Enrollment Report for New York State. United States Department of Education. <https://title2.ed.gov/Public/Report/PrintSection.aspx?Year=2014&StateID=36&Section=140140>
10. Faulkner, S., & Cook, C. (2006). Testing vs. Teaching: The Perceived Impact of Assessment Demands on Middle Grades Instructional Practices. Research in Middle Level Education, 29(7). <http://files.eric.ed.gov/fulltext/EJ804104.pdf>
11. Amrein, A., & Berliner, D. (2002). An Analysis of Some Unintended and Negative Consequences of High-Stakes Testing. <http://nepc.colorado.edu/files/EPSSL0211125EPRU.pdf>
12. McMurrer, J. (2007). Choices, Changes, and Challenges Curriculum and Instruction in the NCLB Era (N. Kober, Ed.). [http://www.cepcdc.org/cfcontent\\_file.cfm?Attachment=McMurrer%5FFullReport%5FCurricAndInstruction%5F072407%2Epdf](http://www.cepcdc.org/cfcontent_file.cfm?Attachment=McMurrer%5FFullReport%5FCurricAndInstruction%5F072407%2Epdf)
13. 2013-2015, Grade 8 only, ELA and math scores combined (294,890 proficient, 995,452 tested). <http://data.nysed.gov/downloads.php>
14. 2011-2013 (156,750 proficient, 418,500 tested). Data source: <http://data.nysed.gov/downloads.php>
15. Based on a proprietary logistic regression model of 2013-2015 ELA and math assessment proficiency vs. the proportion of test-takers in the subgroup of "economically disadvantaged". Data source: <http://data.nysed.gov/downloads.php>