

Andrew Pallotta
President

Jolene T. DiBrango
Executive Vice President

Paul Pecorale
Second Vice President

J. Philippe Abraham
Secretary-Treasurer

July 13, 2018

MaryEllen Elia
Commissioner of Education
New York State Education Department
89 Washington Avenue
Albany, NY 12234

Dear Commissioner Elia,

Now that this year's grade 3-8 state testing is complete and the assessment discussion at the June Regents meeting has taken place, it is time to review where we are with state testing and where we should go to bring trust and credibility back to the system. Because of the lack of trust in the system built up over several years, the tests have been rendered irrelevant to teachers and parents around the state. Simply changing testing companies, reducing the number of days, and providing unlimited time is not enough to restore trust. These changes are not addressing the critical issues of returning our schools to a time when teaching and learning was at the center of everything we do. The New York testing system needs a complete overhaul.

This is highlighted in a recent study from SUNY New Paltz's The Benjamin Center titled, "Tests Are Turning Our Kids Into Zeroes: A Focus On Failing," in which the authors looked at zero scores on Constructed Response Questions (CRQ), including both Short Response and Extended Response questions, from 2012-2016 NYS 3-8 ELA assessments. The trends uncovered in this study show there have been sharp increases in the percentage of zero scores across all grades since the administration of exams aligned to the Common Core. A zero score is given to an answer that is "incomprehensible, incoherent or irrelevant," zeros are not a blank or unanswered question. In particular, children in grades 3 and 4 have struggled with the CRQs. In grade 3 the percentage of zeros jumps from 11 percent in 2012 (pre- Common Core alignment) to 18 percent in 2013 (first year of Common Core alignment) and this increase was sustained through 2016, the last year of this study. The most dramatic rise from pre-Common Core was seen in the grade 4 data where scores of zero went from 5 percent in 2012 to 16 percent in 2015. The authors conclude, "We believe that our findings show that the massive New York State Testing Program was built on poorly developed exams that have returned little educational value and have been particularly problematic for the youngest students." When addressing the change from Pearson to Questar the authors state, "...what has been lacking since the advent of CCLS-aligned testing is a way to hold the testers themselves accountable. It may be that there is a new testing company in town, but without understanding the nature and depth of Pearson's flawed work over five years, we may well repeat those mistakes with Questar."

Our recently released report, “The Tyranny of Testing” is based on this year’s testing experience and shows many students and teachers continue to have negative experiences. Issues regarding computer based testing included poor or unreliable connectivity, as well as individual students struggling with this new mode of testing. With the paper and pencil testing we heard concerns about the age appropriateness of reading passages as well as the length of tests in the early grades. While the tests were reduced from three days to two, the numbers of questions were not reduced at the same ratio. The ELA tests saw a reduction in multiple choice questions from 31 to 24 rather than 20 and a reduction in short response questions from seven to six rather than four or five. In addition, the reading passages are often one or two grade levels above the child’s grade level, and students were expected to read too many passages as part of the exam. The higher number of questions per day led to students spending far more time each day on the tests. No child should sit for hours on end completing these tests. It is simply unacceptable. We must limit time spent testing. The length of the test should be collaboratively agreed upon and should include a maximum. This would assist, for example, students who receive double time per an IEP.

Additionally, it is important to correct some of the information provided to the Board of Regents at the June meeting of the standards and assessment work group. During the presentation, SED staff stated teachers and students were responsible for the exams taking too long to complete. What SED staff should have been pointing out was the length of time spent completing the tests was being caused by the test length being too long for two days, as indicated above by the number of questions and passages on the tests. SED incorrectly blamed students for writing too much and teachers for teaching students appropriate strategies for writing. SED should have acknowledged the decision they made and indicated they planned to further reduce the questions and passages for next year so the tests would be more developmentally appropriate. Instead, they chose to blame students who took the tests and teachers who were teaching students how to be good writers.

The second area of the presentation we must correct is the level of teacher involvement in the test development process. While we acknowledge teacher involvement in all aspects of the process, the presentation made it sound like teachers actually control the development process and that hundreds of teachers are involved. The facts would indicate otherwise. While the groups of teachers participating in item writing are large and diverse, the other committees are too small to properly represent the diversity of the state. The testing company and SED staff tightly control the work of these “committees of teachers” limiting their actual impact. For example, NYSUT’s report on the benchmark setting process demonstrates how SED staff drove the process of setting the cut points on the Common Core tests, setting benchmarks well above other states, and ensuring low proficiency rates on the new tests.

The approach of reducing the number of days of testing and the rush to use computer based testing without making sure the system could handle it, will not restore parents trust in the testing system. In order to do that, the system needs a total re-boot. Here are our recommendations on how to fix the testing system.

We believe the state should change the time of the year when students are taking the tests to May, so that students will have additional instructional time before the assessments.

When it comes to test construction and development, these are our recommendations:

- Include teachers and other stakeholders to determine the purpose of the tests collaboratively. Then, we must limit the application of the tests to that purpose. Educators need to decide what information we want and need from these tests and then build a test to acquire that information.

- Committees of teachers from each grade level that reflect the geographic diversity of the state and the different types of districts such as urban, suburban, and rural should select reading passages from a group of grade appropriate passages and develop all of the questions to be used on the tests. Passages should not be above the grade level being assessed. This approach will ensure passages and questions will match what is actually being taught in the classroom. Teachers would then construct the tests using the questions and passages from these committees. Teachers should decide the difficulty of the tests not the testing company.
- To control the quality of the tests, another large and diverse committee of current classroom teachers must be the final review as a quality control mechanism. The teachers conducting this review should be allowed to make any changes necessary to improve the test.
- When overhauling the benchmark setting process, another group of teachers representing the full range of New York State's children must set the benchmarks and cut scores for the proficiency levels on the tests. This committee will be the key to returning trust and credibility to the testing system. The tests must be benchmarked appropriately, and teachers must determine what it means for a child to perform at each grade level. Much of the problem with the tests deals with the fact that so few people understand what the state means by proficient and the state is using NAEP, which no other state does. The new benchmarks and cut scores must reflect students' progress toward graduation. We must eliminate the gap between performance on the grades 3-8 state tests and the Regents exams and graduation rates. Finally, the criteria for scoring of the tests should be directly related to the items answered correctly.
- Currently some of the companies conducting the grading hire people not certified in the subject being assessed to grade the test. This practice must change. Lastly, the state should ensure that students' responses to the tests are provided to teachers in a format that is instructionally useful.

On computer based testing, SED should:

- Provide transparency by conducting a study of the cost of computer based testing and paper and pencil testing including known and potential costs incurred by school districts to use such testing.
- Review all research based studies on the validity and reliability of computer based testing versus paper and pencil tests to determine the potential impact of the shift to computer based testing on school accountability.
- Answer all of the questions NYSUT previously provided in a letter dated April 30, 2018 to SED staff and the Board of Regents regarding computer based testing (Addendum A).
- Form a committee of practitioners including teachers, principals, superintendents and other educators to review the information requested above and to develop a plan for the appropriate implementation of computer based testing including the grades to be assessed by computer based testing. The plan shall include the cost of providing equal access to the technology needed to participate in the test. The committee make up must be diverse and include practitioners from all over the state.
- Ensure computer based testing is not limiting curriculum instruction. Standardized computer based tests should only assess standards easily assessed on a computer. Teachers should not have to put curriculum on hold to teach computer skills in order for the children to properly take a test on a computer.

As a 25 year veteran teacher, the moments that meant the most to me were those where we were able to effect change by working collaboratively with parents, administrators and the community to make decisions that impacted our students using a shared decision making model. Every committee had

July 13, 2018

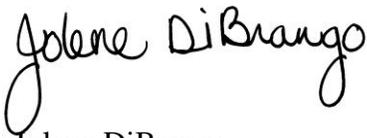
Page 4

rich, diverse representation from the teachers in the classroom, not a sprinkling, but an abundance of teachers helping to influence decisions that impacted their classrooms, profession, and students. We urge SED to utilize a truly collaborative model where practitioners have comprehensive representation on a variety of committees and where these practitioners are chosen by their fellow teachers to be their voice in this process.

Our system needs to make sense to parents, teachers and students if we hope to have them trust the system. Bringing everyone together in this way can do that. If parents trust the system is accurately measuring their child's progress, they will make choices for their children that will reflect that trust.

As always, if you wish to discuss these recommendations further or better yet, if the department would like to implement them, NYSUT stands ready to work with you.

Sincerely,

A handwritten signature in black ink that reads "Jolene DiBrango". The signature is written in a cursive, flowing style.

Jolene DiBrango
Executive Vice President

#106106

c: Members of the New York State Board of Regents

Enclosure