

**Public Hearing
Regents Learning Standards
and High School Graduation Requirements**

**NY State Standing Committee on Education
Rochester, October 8, 2003**

**Testimony submitted by Richard M. Ryan, Ph.D
Educator and Parent**

I testify today not only as a father of two school age children, but also as an educator, concerning the damage the current SED-led test-focused reform is doing to a whole generation of our state's children. I have been in the field of psychology for over two decades, and as an educational researcher I have many ongoing opportunities to work with teachers, administrators and students alike. I also do experimental and field research on the motivational impact of testing, evaluation and teacher strategies.

Given this exposure, I deeply believe that SED's attempt to address the complex needs of NY's diverse learners by trying to force all kids to fit through an identically-sized peg hole is ill-conceived. It is based on neither good common sense nor good scientific research. Common sense tells us that, with their diverse talents, distinct disabilities, varied interests, and unique backgrounds, forcing all of our kids to embrace the same curriculum, jump identical gauntlets, and be assessed with only one tool, does not constitute either an effective approach to education or a legitimate atmosphere for reform. And when many of our kids don't fit through this peg hole-- as the data on escalating dropouts, G.E.Ds and missing students indicates— common sense should tell us why. One size fits all is a regressive approach to schools. Supporting this view, the emerging scientific research tells us that this "high stakes testing" (HST) reforms are producing numerous unwanted and highly deleterious side effects. This research should give us pause. When common sense and science converge, perhaps we should doubly pause.

To complicate this impact, the SED's unreliable, constantly wavering criteria for year to year passage of tests makes the system patently untrustworthy as a sole basis for doling out the high stakes of high school graduation. To make such an important judgement on the basis of criterion that shift from year to year and committee to committee is at least premature, and at worst malpractice.

This is, in part, why professional groups, including the American Educational Research Association, the American Psychological Association among others have spoken out against the misuse of single programs of tests to make high stakes decisions. The current usage of tests as high stakes exams is inappropriate both in terms of its validity and its consequences.

In my testimony today I want to begin with my experience as a father, and how the testing regime damaged a fine public urban school. I then want to move to a summary of research findings congruent with what our family observed: “high stakes” use of tests leads to teaching to the test, less rich learning experiences, more dropouts, less engagement. In the end I provide some references to recent research supporting this view and offer my opinion about the need for rethinking the Regent’s high stakes testing approach to reform.

Our Experience

My family lives in the City of Rochester, and until recently we had a child enrolled in one of the magnet schools within the district, and another planning to. The magnet school was attractive to my daughter because she hoped to pursue some of her interests in the arts during high school, while at the same time receiving a good education.

Pressure on test scores has, however, changed her school. It began well before her 8th grade Regents exams 3 years ago, and extended into high school, dramatically altering the culture of the school for the worse. A nearly myopic focus on test preparation took root in most of her academic classrooms, and creative teaching transformed into an endless stream of worksheets. Teachers became more preoccupied with transmitting isolated facts and teaching “rubrics” for test takers. There was less time for discussion, for critical thinking, or for in-depth projects. Paradoxically, the school’s search for higher test scores was yielding a *lower* quality of instruction. For my daughter it became an increasingly boring and unresponsive place to learn. We saw her begin to lose her love of learning and we couldn’t allow it.

My daughter’s school’s charter was to “draw in” and nourish kids with interests in the arts, combining those from diverse socio-economic and cultural backgrounds. Instead, as a Regents-controlled institution it transformed into a shallower learning environment, and one destined to become less diverse, and less supportive of its diverse constituents.

This occurred in short order after the new Regent’s testing regime went into full swing. As the school focused on test scores, (a focus made explicit in the principal’s speeches and in parent night presentations) the school retreated from its enriched learning themes. Teachers became worksheet oriented, with lesson plans aimed primarily only at the standardized curriculum they anticipated would be tested. My child frequently reiterated comments from teachers about how they were “preparing” for a test, or who used “the test” as the rationale for why a given topic was being taught. Self-directed learning, hands on work, and in-depth writing projects became less and less frequent. The education went from 3 dimensional to just plain linear and dull. It was clear that tests, rather than excellent teaching, had taken front and center.

A few examples: In U.S. history they abandoned rich research projects in which students would learn to research in an in-depth way, instead “moving on” to cover in a superficial survey isolated facts from our nations past. Why? Because that is what would

be tested. In biology they focused on “covering” units in rapid order—abandoning depth to race over all the material that might be tested at years end. In English they spent time learning the State's formulaic approach to essay scoring, as if this “rubric” represented the “right way” to craft an essay. Assigned essays tended to be short, more fitting with what would be required on the target exam.

She no longer goes to that school-- she now attends a private high school that is not subject to the Regents testing regime, and is thus not compromised by it as her prior high school was. As you no doubt aware, it is explicitly because NYS private schools want to maintain their high quality that they, as an association, have not adopted the Regents requirements as a standard. Nonetheless, as firm supporters of City Schools we did not want to have to make that choice for a private school. Despite the hardship, we have enough resources and concern to make that choice for our children. But as a taxpayer, I believe we should have real choices **within** public schools rather than a test-driven “one size fits all” system that offers no diversity or options. My family had to abandon public education to find a better option, and that is wrong. Even more wrong is that so many families do not have that choice at all.

Research Findings

Research findings bear out the kinds of impact our personal experience entailed. Below I offer summaries of new studies examining the impact of High stakes testing reform on our nations children. Please consider these findings from reputable sources (most references below)

1. Research verifying that teaching to the test is a big problem

Moon, Callahan, and Tomlinson (2003) recently surveyed a nationally stratified random sample of teachers on the effects of state High Stakes Test programs on their classroom practices, which included NYS teachers. Results indicated that classroom practices were strongly affected, especially in schools serving students in the lowest socioeconomic strata. Teacher reports suggested increases in test scores are not necessarily a result of student academic attainment but more due to test preparation. **Test preparation was reported to drive out other instructional activities**, as much time was taken in the classroom to review and practice for state testing. Test preparation was especially intense in poorer districts. Moon et al suggested that one result of high stakes testing regimes is a narrowing of the curriculum and the implementation of practices that may actually run counter to effective instruction. Indeed, the authors suggested that the very salience of test scores in the minds of teachers may be **restricting educational opportunities, particularly among those from the most impoverished areas**. Moon et al. further suggested that when teachers specifically teach to the test, the scores may no longer represent the broader domain of knowledge for which they are supposed to be an indicator, especially in schools serving disadvantaged students where the test prep was reported to be more intensive.

A study by **McNeil and Valenzuela (2000)** arrived at similar conclusions in their study of Texas teachers. They found that teachers were encouraged or required to reallocate time away from core subjects not tested on the state exams, and to eliminate or curtail special projects, experiments, library research, extensive writing, or oral assignments. This was especially true in schools that might be lower in absolute performance levels (i.e., those serving less affluent students). Much time was also reported being spent specifically on test taking strategies rather than substantive issues.

2. Failure of test score gains to transfer to other indexes of achievement.

Evidence that HSTs lead to “teaching to the test” – which in turn crowds out the teaching of skills not on the tests and the provision of enriched experiences that might better engage student interest in additional knowledge seeking – may underlie the concern with the generalizability of score gains. Researchers use the word *transfer* to discuss this—transfer concerns the extent to which gains on any given test are reflected in improved achievement on other, non-targeted measures. Little research exists on the validity of test score increases on HSTs, despite the fact that it is a crucial bone of contention between HST (High Stakes Test) advocates and their opponents.

Perhaps the most comprehensive research examining this issue was an 18-state study by **Amrein and Berliner (2002)**. To test the transfer of score increases on high stakes exams they obtained scores on non-HST tests that overlap with HSTs in their assessment of achievement domains. These were the ACT, SAT, NAEP and AP tests. Their evidence suggested, contrary to HST advocates, that when transfer is considered, **level of learning in those states with salient HST policies remains level or falls below previous levels once HSTs are implemented.** In contrast, states without high stakes graduation tests were more likely to show improvements on these outside tests than states that had imposed them. Indeed more than two-thirds of states posted decreases on ACT performance after high stakes graduation exams were implemented.

Neil & Gaylor (2001), using the NEAP as a metric, similarly showed that states without HSTs were more likely to show score improvements than states with them. That is, NEAP scores were not improved by HST initiatives, and they also carried many other potentially negative consequences with them. They specifically suggested that **HSTs widen educational outcome inequities between rich and poor rather than ameliorate them.**

3. Dropouts as a direct result of the Misuse of Tests.

With so much attention paid to state test scores because of the SED pressure, an equally important gauge of school performance often gets ignored—namely, high school **dropout rate**. Although dropouts are hard to track (and are often systematically misreported) available data show that both dropouts and students leaving high schools for equivalency diplomas are on the rise, with notable escalation in the past few years as HST policies have intensified.

Reardon and Galindo (2002), for example, studying students across the nation between eighth and tenth grade in districts with and without High Stakes Testing policies, estimated that the imposition of HSTs **increased the odds of dropout by 39 percent**.

Although accounts differ, one possibility is that as states required students to pass tests for promotion, more pupils were held back. This has been verified in NYS where retention has gone way up in the 9th grade, prior to the testing years for the Regents subject area exams. In turn, well-established research shows that the mere fact of retention dramatically increases the probability of dropout. In addition, if one assumes that HSTs impose even modestly more difficult standards, that too could lead to amotivation and discouragement among students already at risk for failure. When disadvantaged and/or low achieving students are forced or encouraged to drop out, this may make a school's average scores look better, even while society as a whole pays the price.

A related issue is the concern that HSTs may lead many to seek **equivalency diplomas** (G.E.D.). Studies comparing high school graduates to young people who received equivalency diplomas show that even among those with similar academic scores, those who complete high school have higher earnings, better employment, and commit fewer crimes. One reasonable account of this is that the confidence, self-esteem and work habits of young adults is greater if they graduate from high school than if they drop out to earn a G.E.D., and that confidence translates into better adult outcomes. In other words, if HSTs drive students out of school this has costs, most of which will be borne by children from lower income families.

Jacob (2001) examined the effects of high stakes high school examinations on student retention especially among low achievers. His findings, based on analysis of data from 15 states, showed **that students in the bottom 20th percentile of achievement who faced such requirements were 25 percent more likely to drop out in states with tests**. He also found, however, that use of the tests had no significant effect on subsequent academic achievement for the population considered as a whole.

The dropout pattern in New York State is already clear, even though the HSTs will not be fully phased in until 2005. The NYS Education Department's own statistics (2002) reveal an increasing dropout rate. Particularly affected are students for whom English is a second language, whose dropout rates nearly doubled.

4. The unreliability of passing rates: Can Regents Scores be used to gage reform?

Tests inform best when they are used to gain information about where educational programs are effective, and when they are reliable gages of progress from measurement point to measurement point. However, NYS tests, which are admirably reliable within each test (SED data shows, e.g., good internal consistency), have shown themselves to often be unreliable and to be based on wavering standards for passing between one test

year and the next. This reflects the fact that there is no clearly established validity behind the established cutoff scores or criterion for passing a test on any given year.

In fact, the procedures for establishing these criteria in NYS remain mired in mystery and guesswork. Some years passing rates on specific tests have been wildly high, whereas the “same” tests the following year will be low. Clearly, passing rates on the same subject matter have regularly fluctuated as much as 20-30% points over a single year. Is this telling us about year-to-year changes in schools, or does it suggest unreliable passing criterion? One need only look at the recent Regents Physics exam fiasco or the 2003 Math A exam, to see concrete examples. Furthermore, there is little evidence to suggest that when a given test improves across years that this is not because the test itself becomes easier, rather than because students have, on average, gained more skills. Scores can go up because tests change in difficulty, because passing scores are adjusted, or because teachers more and more teach to the test. All this practices make the meaning of test score changes more dubious, and interpreting them naively can yield inappropriate conclusions about effects on diverse groups. For example making the 4th grade ELA test easier might give the appearance of closing an achievement gap between groups, when indeed it really just increases percent passing rates without reflecting skill changes. In addition, year-to-year passing rates do not take into account changes in the test pool due to increased “pushouts”, increased dropouts, increased G.E.D. enrollments, or increased classification of children in IEP’s that exempt them. As all of these categories increase, an artifactual increase in average test scores of those who remain can emerge.

The lack of a standard criteria or anchor applied year to year thus makes **test “passing rates” an unreliable gage of progress in reform**, and at the same time puts many student’s future up for grabs. Under the current NYS testing scheme, the standards of graduation for a given child are more arbitrary than ever. For low achieving kids, they may find themselves below standards on a given year when a Regent’s panel decides to “get tough” on math A. Or he/she may find scores above threshold when a panel eases its requirements because of a poor showing on a previous exam (e.g., note the change in Physics passing rates since last year.)

Testing Without the High Stakes

I want to finish by saying that I am not against testing. Collecting standardized testing information through appropriate sampling methods (such as those currently used in Germany) could be a very important way to help schools be informed about practices that work, and areas in need of improvement. Testing allows accountability, but when bad policy primarily tests as a high stakes whipping device, a purpose for which tests were not created, bad practices result.

The mounting evidence from outside SED documents that the “reform” strategies being employed by the Regents under Richard Mills are leading to more dropouts, to excessive and counterproductive teaching to the test, and to an unresponsive learning environment for the vast majority of NYS’s students. In addition the tests are flawed in their construction and implementation, and used for purposes for which such tests were

not intended. We need more flexibility, more choice, more information, and less rigidity from SED.

Citations:

Amrein, A.L. & Berliner, D.C. (2002) High stakes testing, uncertainty and student learning. *Education Policy Analysis Archives*, 10(18). Retrieved from <http://epaa.asu.edu/epaa/v10n18/>.

Jacob, B. (2001). Getting tough? The impact of high school graduation exams. *Educational Evaluation and Policy Analysis*, 23, 99-121.

McNeil, L., & Valenzuela, A. (2000). *The harmful impact of the TAAS system of testing in Texas: Beneath the accountability rhetoric*. Cambridge, MA: Harvard University Civil Rights Project.

Moon, Callahan & Tomlinson: Effects of State Testing Programs on Elementary Schools with High Concentrations of Student Poverty—Good News or Bad News? *CIE* V6 N8

Reardon, S.F. & Galindo, C. (April, 2002). *Do high stakes tests affect students' decisions to drop out of schools? Evidence from NELS*. American educational research Association, New Orleans.

Grissmer, D., Flanagan, A., Kawata, J., & Williamson, S. (2000). *Improving student achievement: What state NAEP test scores tell us*. Santa Monica, CA: RAND, MR-924-EDU.