

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

**NY State Standing Committee on Education  
NYC, October 15, 2003**

**Testimony submitted by Robert Stake  
Professor, University of Illinois**

---

I am Robert Stake, Professor emeritus of Education, Director of the Center for Instructional Research and Curriculum Evaluation at the University of Illinois. I am a theoretician and practitioner of educational evaluation. I was trained in Psychometrics at Princeton University and was a Psychometric Fellow at the Educational Testing Service. I have been Associate Director of the Illinois State Testing Program. I have 40 years of experience evaluating schools and national programs, including ten years as an external evaluator of teacher staff development in the Chicago Public Schools. My evaluation designs put a major emphasis on classroom observations and interviews and a lesser emphasis on testing and other instrumentation. I testified in 1991 before a panel of the New York State Board of Regents. (My travel expenses for this trip were paid by the Consortium.)

I am pleased to have the opportunity to present research-based findings regarding the use of testing as the primary basis of determining student readiness for graduation.

I was a member of the Blue Ribbon Panel appointed in 2000 by Commissioner Richard Mills to evaluate the request of the Consortium for Performance Based Assessment for variance from the use of Regents Examinations for graduation. The other five members of the panel were test-oriented researchers who had previously served Commissioner Mills as consultants on the development of the new Regents examinations.

The Panel report was submitted in 2001. The Commissioner concluded his testimony today with the words, "Children in New York are learning more now than a decade ago. Testing is one part of the system that made that happen." There is no credible body of evidence to support the claim. It is more likely that the quality of education for most children has dropped with teachers focussing narrowly on test content and with increased numbers of students poorly motivated and dropping out. Testing is probably the main cause of lowered quality of teaching and learning.

In his remarks to the Senate committee, the Commissioner referred to our Blue Ribbon Panel. One could easily misunderstand what he said. Allow me briefly to summarize our report. It is important.

Our report pointed out that the Consortium Schools were asked to meet a higher standard for assessment than were the developers of the Regents examinations. The panel also found that voluminous records provided by Consortium schools had little evidentiary quality but that the definition of evidence set forth by the State Education Department was inappropriate for an approach that integrated performance assessment with individualized learning. The Panel found articulation of State Standards inadequately connected (at the lesson level) with teaching and learning at the Consortium schools but panel members relied on a psychometric impression of alignment rather than drawing upon the relevant expertise of curriculum specialists. When the Panel found evidence of high satisfaction of students and parents, low drop-out rates, and unusual success in college after graduation, it concluded that these good outcomes probably were in part attributable to Consortium's unconventional assessment procedures. The Panel recommended that the variance be continued for three years and that independent evaluation and research be done to assure that equal and fair learning opportunities were not being sacrificed.

Commissioner Mills acted contrary to the findings of the panel, a panel made up primarily of his own consultants.

Let us shift to the broader picture of school improvement.

1. The strategy of educational accountability and school improvement by mandated standardized student testing has been tried in almost all states of the union, in Michigan, Florida and California for more than a quarter of a century. This strategy has not successfully raised the quality of teaching and learning in schools in the nation or slowed the deterioration.
  - a. The nation and New York State are without a respected body of evidence to support claims that teachers now are more effectively teaching mathematics and language arts and other academic subjects than they have in the past.
  - b. There has been an enduring politically-powerful movement to establish state standards and high stakes testing in New York and other states. It is clear that this has resulted in “more uniform” instruction. For some children this has helped make the teaching better. For more children, probably for a majority of children, this standardization has made the teaching more simplistic, less personally meaningful, less motivating and more intimidating. Testing has contributed to watering down the education of most youth and to driving thousands of students out of school before graduation. (See Haney testimony. Sept 23)
  - c. Some educational philosophers such as John Dewey advised us that you should look for quality in education by examining the experience that learners are having more than by looking at their performances in artificial situations such as testing. One of Dewey’s students, Ralph Tyler, was the creator of NAEP, the National Assessment of Educational Progress. It is clear from his writing that Tyler would not have approved requiring students to pass the Regents Examination as a requirement for graduation. He was not opposed to controls over the teaching profession but he opposed using testing for that purpose.
  - d. Making testing a requirement for graduation actually makes it an ineffective measure of achievement. Judging the readiness of a student to leave high school should be based on the review of many student performances and the deliberation of many professional people, with contributions to the decision by the individual student. Society, higher education, and the workplace are not served by determining eligibility for graduation heavily weighted by testing.
  - e. The ineffectiveness of the public education we have today cannot be fixed by making better tests and scoring techniques. It is a much bigger problem. I doubt that education can be fixed by chief state school officers who, in effect, have only one strategy, student assessment. The best model for state control of schooling today is in Nebraska, where a balance between state, local and teacher control is vigorously pursued. There, professional educators and state officials are not adversaries in the way they are in New York. (Douglas Christiansen is the commissioner.)
2. The new Regents examination items clearly have subject matter content. But getting the items right is a matter of reasoning, understanding how to take tests, and clever guessing more than a matter of knowing the knowledge being tested for. The tests encourage students and teachers to concentrate on how tests work more than on getting comprehensive and experience-based understanding the subject matter.

- a. Please consider this test item: Who said, “I am fighting this war so that my son may plant grain and his son may paint pictures.” George Washington, John Adams, Dwight Eisenhower, or Ronald Reagan. A few students will have learned that it was John Adams. Most who get it right eliminate Eisenhower and Reagan as too modern to talk like that and they won’t remember anything about Washington connected to farming and art. And they will chose Adams partly because he is the most obscure. Yes, good guessing requires some knowledge, but this test item is not testing what the test users intended.
  - b. The name for what the tests are testing has been scholastic aptitude, a competency common across subject matters. Test makers shy away from the term today because it suggests they are testing innate intelligence. They are testing for culturally developed intelligence. There is a problem in claiming that the tests are fair across cultural backgrounds.
  - c. The item pools today in state departments across the country are measuring pretty much what aptitude tests measured before and after World War II. It is aptitude more than achievement. Teachers cannot increase student scores much by teaching superior course content, but they can directly increase scores by raising test-taking skills. (Professor Robert Meally and the Language Arts Panel spoke with disgust about such mismeasurement.)
  - d. The views that teachers have of student competence is different from the views that test makers have. A pedagogic view concentrates on how students perform in episodes in class or elsewhere. A psychometric view concentrates on traits, intellectual descriptors. Traits and episodes do not easily transform one into the other. Both are useful for examining student progress. A concentration on traits alone oversimplifies the representation of progress toward becoming well educated. (My paper on Pedagogical and Psychometric Perception is available.)
3. Standardization contributes greatly to the problem. The definition of fair and equal education does not require the same subject matter content, the same skills, the same passing scores for everyone. Opportunity and culture should influence the aspired-to education of individual students. It is more important that schools optimize learning experiences than to assure that everyone learns a minimum amount. Standardization is important in car parts and pharmaceuticals but not in educating our youth.
    - a. Using a court decision to justify standardization of student performance is a political interpretation rather than an educational interpretation of justice. A diploma has never meant the same thing and should not mean the same thing in all the schools of the state.
    - b. Graduation readiness should be decided primarily by professional people who have personally worked with the student. It is a human subjectivity. They should consider standardized test scores, portfolios, and other evidence of taking advantage of opportunities for gaining knowledge, career readiness, and social responsibility.