
Keeping Accountability Systems Accountable

It is widely assumed that if students pass state tests that are aligned with state standards, their school is doing a good job of educating them, and if they pass the state's exit exam, they are ready to face the challenges of college. Ms. Foote questions those assumptions and presents an alternative way of judging the quality of a student's high school education.

BY MARTHA FOOTE

THE STANDARDS and accountability movement in education has undeniably transformed schooling throughout the United States. Even before President Bush signed the No Child Left Behind (NCLB) Act into law in January 2002, mandating annual public school testing in English and math for grades 3-8 and once in high school, most states had already instituted their own accountability systems of state standards linked to state exams. Throughout the country, a battery of tests — the FCAT in Florida, the TAAS and later TAKS in Texas, the SOL in Virginia, the MCAS in Massachusetts, and the AIMS in Arizona, to name a few — awaited students in a few select grades each year. States used the

MARTHA FOOTE is the director of research for the New York Performance Standards Consortium, New York, N.Y.

results of those tests to determine their educational health and, in turn, to judge their schools. Now, with NCLB, the number of students tested annually has skyrocketed as all 50 states have exams in operation, with even more grades — seven in all — required to administer tests.

Though critics have denounced so-called high-stakes testing for reducing curricula to circumscribed test content and learning to rote memorization,¹ proponents have countered that the standards and accountability movement has spurred increased learning, as shown by rising state test scores.² Many states have reported test score gains, especially in the early grades, but questions are emerging as to whether these increased test scores really do indicate increased learning — or, for that matter, if the tests measure meaningful learning to begin with. Because these state tests are the accountability

measure for schools and students, bringing potential real-life consequences to them both, these questions must be asked and answered.

STANDARDS AND ACCOUNTABILITY

The basic premise behind “standards and accountability” is quite simple. First, states determine which content and skills — the “standards” — their students need to learn. Teachers then teach to these standards, and state tests measure whether students have indeed met them. The various players are then held “accountable” for the results as a wide range of punishments and rewards kicks in. Under NCLB, if schools do not meet their state’s Adequate Yearly Progress (AYP) goals, as measured by gains in student test scores, they face increasingly punitive actions that can culminate in school restructuring and state takeover, potentially leading to job loss for teachers and principals.³

As for rewards, NCLB promises public recognition of high-achieving schools as well as monetary bonuses to teachers in schools that make the greatest gains as measured by test scores.⁴ Similarly, policy makers in some states provide their own sets of financial incentives to teachers and principals, as well as public acknowledgments of achievement. Then there are always the newspaper headlines, routinely praising or shaming neighborhood schools on the basis of test results. In short, accountability becomes synonymous with a public display of judgment.

Many states have also zeroed in on students as targets of their accountability systems. According to the Center on Education Policy, roughly half of the 50 states have enacted, or are in the process of enacting, a mandatory high school graduation test.⁵ By requiring such a test, state policy makers demand accountability from students to show that they have met the state standards. Proponents of these high-stakes tests assert that students will work harder, and thus learn more, when they know that their diploma depends on passing such a test.⁶ Critics, however, point to a link between state graduation exams and higher dropout rates for lower-achieving students, African Americans, and Latinos, as well as to an increased incentive for schools to “push out” teens who are academically at risk, effectively limiting these students’ futures.⁷

Finally, hovering somewhere above this sharp focus on schools, principals, teachers, and students are the policy makers responsible for mandating education policy in the first place. Though not a direct variable in the

standards and accountability equation, policy makers can be held accountable for the decisions they make. Citizens can vote their politicians out of office. However, state education policy is often decided not by the politicians themselves, but by their appointees. In New York, for example, the state legislative assembly selects the 16 members of the Board of Regents, and the Regents, in turn, select the state education commissioner. The commissioner sets policy, and the Regents officially approve it, with no elected official directly accountable to the voting public.

Suspiciously absent from the accountability equation are the tests themselves. Though these exams are the measure by which schools, principals, teachers, students, and even policy makers are publicly deemed successes or failures, many states refuse to release them for public scrutiny. Even in those states where the tests are made available, they are often left unexamined, shielded in essence by an a priori assumption that they are a gauge of a good, solid education that will pave the way for students to achieve academically at even higher levels. But is that assumption really valid? Recent results on the National Assessment of Educational Progress (NAEP), a nationwide reading and mathematics exam, have certainly called this belief into question: despite soaring state test scores following the mandatory implementation of state standards and tests as required by NCLB, trends in the NAEP reading test have stagnated for fourth-graders and declined for eighth-graders, while slowing in growth for both fourth- and eighth-graders in math.⁸ Furthermore, in many states requiring exit exams, a growing body of evidence suggests that high school graduates (i.e., those who *passed* the exams) are unprepared for the academic challenges of college, despite the insistence of officials that, with the test, the diploma means real achievement. For example, since Texas instituted its graduation exam, it has seen a rise in the percentage of its high school graduates requiring remedial classes at public colleges or universities,⁹ while Massachusetts public colleges continue to see the same rates of unacceptably low skill levels in their entering students as they saw before the state’s graduation exam went into effect.¹⁰

Though state tests purport to measure whether a student has met state standards, just how meaningful is passage of a state test if that achievement does not translate into deeper learning and subsequent academic success? In other words, which standards are actually being met and measured, those that require simple memorization or those that demand complex thinking?

Isn't it time that we hold these tests accountable for what their proponents claim they can do, especially when children's futures and adult livelihoods hinge on the results? At the New York Performance Standards Consortium, we believe the answer is yes.

HOLDING THE ASSESSMENTS ACCOUNTABLE

Public accountability and transparency have been crucial to the success of the New York Performance Standards Consortium, a coalition of 28 small, diverse public high schools across New York State that exemplify education reform based on a strong commitment to school-as-community, to ongoing professional development, and to innovative curricula and teaching strategies. Recognizing that their students learn best when actively engaged, consortium schools typically use inquiry-based methods of learning with classrooms steeped in discussion, project-based assignments, and student choice. Consortium schools are also committed to using complex, performance-based assessments to gauge student learning, with four specific performance tasks required of all students for graduation — an analytic literary essay, a social studies research paper, an original science experiment, and the application of higher-level mathematics. These assessment tasks, which are graded with detailed rubrics by teachers and, through an additional layer of accountability, by external evaluators, constitute a major portion of the consortium's assessment system.

This system is regularly overseen by a group of educational experts known as the Performance Assessment Review (PAR) Board, whose charge is to ensure the quality of student work across schools as well as the efficacy of the consortium's assessment system within each school. The PAR Board, then, holds the consortium accountable for maintaining a high level of student and school performance.

To show that its schools successfully engage and educate their students, the consortium has publicly disseminated an array of statistics comparing its schools in aggregate to New York City high schools in general. Consortium schools post a lower dropout rate, higher college-bound rate, and higher daily attendance.¹¹ These data are especially impressive because consortium schools have more students of color, more students who qualify for free or reduced-price lunch, more students receiving special education services, and more entering ninth- and 10th-grade students scoring below the state standard on reading and mathematics tests than the aver-

age New York City high school.¹²

The consortium, however, has not been content to rest on its accountability laurels in the area of high school statistics alone. Eager to confirm that its approach to education truly prepares students for the rigors of college-level work (i.e., to determine the system's predictive validity for academic success in college), the consortium began work in 2001 on a longitudinal study of the college performance of its schools' graduates. In contrast with testing proponents who claim that graduation exams in and of themselves are valid markers of academic achievement, the consortium has turned to the actual college grade-point averages (GPAs) and persistence rates of its graduates to help evaluate whether its system of teaching and performance-based assessments provides a solid foundation for the academic demands of college. The results, discussed below, are currently available for the high school classes of 2001 and 2002.

Notably, the consortium does not use state test scores as a measure of accountability because its schools are exempt from New York State's graduation exams. In 1995, Thomas Sobol, New York's former commissioner of education, granted the consortium schools a waiver from the state's Regents exams. He thus supported them in their endeavors to develop and utilize a transparent, externally validated performance-based assessment system — reliable across schools — that would ensure complex teaching and in-depth learning up to the state standards and beyond. Several years later the Board of Regents, guided by Sobol's successor, Richard Mills, instituted a new state policy of high-stakes testing and

deemed passage of five Regents exams mandatory for graduation. Despite the policy change, however, the consortium's waiver has been permitted to continue, allowing students to substitute the consortium's performance assessments for the Regents exams, largely because of the success of the schools.¹³

Results from the study have been impressive.

In the sample, 77% of consortium school graduates attended four-year colleges, 19% attended two-year colleges, and 4% attended vocational or technical programs.

No comparable study has been attempted by the state to determine whether the Regents system of five high-stakes tests adequately prepares students for college. To date, its predictive validity for college success remains unknown. Though state officials proclaim that New York's graduation test policy increases the achievement levels of students, they have publicly provided no evidence beyond rising Regents test scores.¹⁴

THE COLLEGE PERFORMANCE STUDY

Each year, all 28 schools in the consortium are invited to participate in its college performance study. Those that join the study ask their graduating seniors to sign release forms granting the consortium's research director permission to obtain their college transcripts. In June 2001, 18 of the 28 schools agreed to participate. In June 2002, 15 consortium schools — including 13 from the previous year — agreed, for an overall school participation rate of 59% over the study's first two years. The average participation rate for each school's seniors is 74%. The schools' population for the entire sample is 20.1% white, 26.0% black, 44.0% Hispanic, 9.8% Asian and others; 60.3% are eligible for free or reduced-price lunches. These numbers approximate those for the entire consortium.

The first round of signed release forms was sent to colleges in December 2002. The second round was sent in February 2004. Of the 967 senior release forms sent to colleges, 666 resulted in transcripts, for a return rate of 69%. Colleges that did not return transcripts overwhelmingly gave one of three reasons: no enrollment record (due to inevitable changes in college plans

as students completed the release forms while still in high school), insufficient data on the senior release form to locate the student in the college database, or a hold on the transcript for internal institutional reasons. Only one college refused outright to honor the senior release form.

The 666 transcripts were analyzed and coded for GPAs, number of credits counted toward GPAs, certain descriptors (i.e., college governance and selectivity),¹⁵ and current enrollment status. To increase reliability, individual GPAs were weighted against the number of credits accrued. Statistical analyses were then run to obtain the percentages of students attending four-year colleges, two-year colleges, and vocational

programs; four-year college selectivity enrollment rates; average GPAs; and the persistence rates of those students who had entered college within one year of high school graduation.

Results from the study have been impressive. In the sample, 77% of consortium school graduates attended four-year colleges, 19% attended two-year colleges, and 4% attended vocational or technical programs. In the sample of students attending four-year colleges, 7% enrolled in the most competitive colleges, 14% enrolled in highly competitive colleges, 30% enrolled in very competitive colleges, 32% enrolled in competitive colleges, 14% enrolled in less competitive colleges, 2% enrolled in noncompetitive colleges, and 1% enrolled in specialized colleges.¹⁶

Upon completion of up to three semesters of college, the average GPA for consortium schools' graduates in the sample was 2.6 out of 4.0, which is approximately a B-. For students attending four-year colleges, the average GPA was 2.7. For students attending two-year institutions, the average GPA was 2.2. It is worth noting that the ACT defines college readiness as the ability to earn at least a C, or a 2.0 GPA, in college-level courses.¹⁷

Consortium students tend to persist in college as well. Of those in the sample who entered college within one year of high school graduation, 78% overall enrolled for a second year. Of those attending four-year colleges, 84% enrolled for a second year. Of those attending two-year institutions, 59% enrolled for a second year. In comparison, nationally only 73% of students who enter four-year colleges and 56% of those who enter two-year institutions return for their second year.¹⁸

These results, combined with the consortium's high school statistics, indicate that consortium schools are highly effective. They hold onto their students, teach them well, graduate them, and send them on to higher education prepared to accomplish college-level work and persist in their studies. When we keep in mind that these schools serve an even more disadvantaged student population than do New York City high schools in general, the outcomes are especially striking. Using the consortium's system of teaching and performance-based assessments, these high schools have found a way to help students succeed in school and beyond.

REAL ACCOUNTABILITY

In the zeal to improve our public schools, state tests have become the keystone of both our national and state educational accountability systems without the necessary oversight with regard to their substance, depth, and validity. A single test score can prevent a student from receiving a high school diploma and building a viable future, yet the test itself may be a meaningless measure of present achievement or future performance. Aggregated state test scores can tag a school as successful, yet its graduates may perform poorly in college. How can we as a nation continue to accept the significance of these tests with only blind faith and the proclamations of our politicians and the test designers as assurance?

The New York Performance Standards Consortium has shown that real accountability centers on more than a mere test score. High standards are met when students accomplish time-intensive, in-depth work that requires complex thinking and analytical skills. Real accountability is achieved when an assessment system demands excellence not only from students, teachers, and principals, but also from itself, with an oversight mechanism for external validation as well as ample evidence of student success beyond test scores. Finally, real accountability must include proof of an assessment system's predictive validity with hard data that correlate the passage of specific assessments with subsequent performance in school, college, or the work force. We must begin to hold our accountability systems accountable, for it is only then that we can truly determine the soundness of student achievement, the quality of our teaching, and the depth of learning in our schools.

1. Alfie Kohn, *The Case Against Standardized Testing: Raising the Scores, Ruining the Schools* (Portsmouth, N.H.: Heinemann, 2000); and Susan

Ohanian, *One Size Fits Few: The Folly of Educational Standards* (Portsmouth, N.H.: Heinemann, 1999). I am indebted to Linda Darling-Hammond and Jacqueline Ancess for their ground-breaking studies of the schools in the Julia Richman Education Complex and to Michelle Fine for her continued research on small schools.

2. S. Anne Hancock, "There's No Doubt No Child Law Works," *Atlanta Journal-Constitution*, 4 August 2004, p. A-13.
3. Erik W. Robelen, "An ESEA Primer," *Education Week*, 9 January 2002, pp. 28-29.
4. Sharon L. Nichols, Gene V. Glass, and David C. Berliner, "High-Stakes Testing and Student Achievement: Problems for the No Child Left Behind Act," September 2005, p. 7, available at <http://edpolicylab.org>.
5. Keith Gayler et al., *State High School Exit Exams: A Maturing Reform* (Washington, D.C.: Center on Education Policy, August 2004), p. 5, available at www.cep-dc.org.
6. *K-12 Education Reform* (Washington, D.C.: Business Roundtable, February 2001), pp. 8-9.
7. Brian A. Jacob, "Getting Tough? The Impact of High School Graduation Exams," *Educational Evaluation and Policy Analysis*, Summer 2001, pp. 99-121; Walt Haney, "The Myth of the Texas Miracle in Education," *Education Policy Analysis Archives*, 19 August 2000, available at <http://epaa.asu.edu/epaa/v8n41>; and *Pushing Out At-Risk Students: An Analysis of High School Discharge Figures* (New York: Public Advocate of the City of New York, 21 November 2002).
8. Sam Dillon, "Education Law Gets First Test in U.S. Schools," *New York Times*, 20 October 2005, p. A-24.
9. Haney, op. cit.
10. Maria Sacchetti, "Colleges Question MCAS Success," *Boston Globe*, 26 June 2005, available at www.boston.com.
11. New York City Department of Education, "Annual School Report Cards" (2002-03) and New York State Department of Education, "New York State School Report Card Comprehensive Information Report" (2002-03). The data for consortium schools were compiled from each school's individual report card or information report. The overall New York City high schools' data are found on every New York City school report card for comparison purposes. However, because the "Annual School Report Cards" base college-bound rates on student reports of intention, which are historically unreliable, the consortium schools' college-bound rate was instead calculated from school reports on the class of 2004, prepared for the New York Performance Standards Consortium in August/September 2004.
12. Ibid. For another example of this type of high accountability, see Michelle Fine, Brett Stoudt, and Valerie Futch, *The International Network for Public Schools: A Quantitative and Qualitative Cohort Analysis of Graduation and Dropout Rates—Teaching and Learning in a Transcultural Academic Environment* (New York: Graduate Center, City University of New York, June 2005), available at www.internationalsnps.org.
13. The only exception is the English Language Arts Regents exam, the one state test that Commissioner Mills ruled was not exempted by the original waiver.
14. Benjamin Wachs, "Putting the Test to the Test," *Brighton-Pittsford Post*, 28 April 2004, p. 1.
15. As determined by *Barron's Profiles of American Colleges: 25th Edition, 2003* (Hauppauge, N.Y.: Barron's Educational Series, 2002).
16. Ibid.
17. Tamar Lewin, "High School Seniors Weak in Math and Science Tests," *New York Times*, 20 August 2003, p. B-8.
18. Thomas Mortenson, "Freshman-to-Sophomore Persistence Rates by Institutional Control, Academic Selectivity and Degree Level, 1983 to 1998," *Postsecondary Education Opportunity*, research newsletter, Mortenson Research Seminar on Public Policy Analysis of Opportunity for Postsecondary Education, Oskaloosa, Ia., August 1998. ■

File Name and Bibliographic Information

k0701foo.pdf

Martha Foote, Keeping Accountability Systems Accountable, Phi Delta Kappan, Vol. 88, No. 05, January 2007, pp. 359-363.

Copyright Notice

Phi Delta Kappa International, Inc., holds copyright to this article, which may be reproduced or otherwise used only in accordance with U.S. law governing fair use. MULTIPLE copies, in print and electronic formats, may not be made or distributed without express permission from Phi Delta Kappa International, Inc. All rights reserved.

Note that photographs, artwork, advertising, and other elements to which Phi Delta Kappa does not hold copyright may have been removed from these pages.

Please fax permission requests to the attention of KAPPAN Permissions Editor at 812/339-0018 or e-mail permission requests to kappan@pdkintl.org.

For further information, contact:

Phi Delta Kappa International, Inc.
408 N. Union St.
P.O. Box 789
Bloomington, Indiana 47402-0789
812/339-1156 Phone
800/766-1156 Tollfree
812/339-0018 Fax

<http://www.pdkintl.org>