

The Abell Report

What we think about, and what we'd like you to think about

Published as a community service by The Abell Foundation

Who's and What's at Stake in "High-Stakes Testing?"

ABELL SALUTES: The Violence Intervention Program at U of M's Shock Trauma: Saving victims' lives, taxpayers' money

It is an all-too-familiar picture: the Shock Trauma helicopter winging ominously overhead, and the sense that some one is badly hurt and is being helped, and, in many cases, a life is being saved. But the picture cannot show that the victim is very often a repeat victim—that he or she is being treated by Shock Trauma for the second or perhaps a third time, and at a cost to the community in each instance of another \$42,000.

In 1998 Drs. Paul Stolley and Carnell Cooper completed a study that examined the horrific waste both in human terms and in dollars of this recidivism. Findings from the study prompted the leadership of Shock Trauma, with support from the Governor's Office of Crime Control and Prevention and Friends Research Group, to create the Violence Intervention Project (VIP). The Abell Foundation contributed start-up and ongoing operating funds to the VIP, which has a twofold mission: to structure an intervention designed to reduce the recidivism and to create a sys-

continued on page 12

Maryland is at a critical juncture in its deliberations on the implementation of the new high-stakes High School Assessment program. The State's decisions will determine who graduates from our high schools, and with what qualifications.

History, background, problems, recommendations:

The Stakes

This June, nearly half of all public school students in the United States were required to pass one or more statewide exit exams in order to graduate from high school. In the next six months, the Maryland State Board of Education (the Board) will similarly decide whether to raise the stakes by making passing scores on the new more rigorous High School Assessments (HSAs) a condition for graduation. In cases like this, where diplomas are contingent on the results of such assessments, tests are characterized as "high-stakes" for students.

If the experience of other states is any indication, Maryland's decision could result in as many as 20% of high school students in the Class of 2008 being barred from graduation as a result of failing one or more of these tests. It will come as no surprise that those denied diplomas would most likely and disproportionately be disadvantaged students from Baltimore City and Prince George's County or students who receive special education services.

Decisions on implementing the HSA program will have irreparable consequences for Maryland's children, yet neither the public nor educators have been privy to much of the dialogue or data sur-

rounding such a critical move. Moreover, the environment in which these decisions will be made is very different, and more formidable, than a decade ago when the HSA program was first conceived. The Board's task is challenging: how will it raise standards for all students without unduly penalizing those who are disadvantaged?

The Challenges

Maryland's HSA program is a series of high school end-of-course exams designed to hold students accountable for mastering specific skills and content knowledge. In addition to holding students, teachers, schools, and districts accountable for meeting clear academic standards, the State Board created the High School Assessment program to:

- increase academic rigor;
- enhance the value of the Maryland high school diploma; and
- ensure that graduates have skills and content knowledge.¹

Whether passing the HSAs will indeed become a graduation requirement in Maryland remains up in the air, but as now proposed, students entering high

continued on page 2

school in Fall 2004 would have to pass at least four HSAs to graduate from high school in 2008.

Maryland's HSAs reflect a national trend to make students demonstrate mastery of specific skills or content knowledge by taking high-stakes exit exams. The trend was accelerated by the federal No Child Left Behind Act of 2001 (NCLB), which requires all states to adopt new high school tests to measure school, state and district accountability. But as states pursue this route, they are confronting tough challenges compelling them to rethink and delay implementation of the testing. Maryland is no exception.

Educational concerns and financial uncertainties have already forced the Board to delay making the HSAs a graduation requirement three times, most recently in May 2003. Meanwhile, students are taking the tests and their scores are being reported on their transcripts as percentile rankings, yet no proficiency levels to measure whether they have met course standards have been set.

This summer, the Board will begin making critical decisions about the future of high-stakes high school testing in Maryland. The Board's charge boils down to addressing a few key issues and questions:

- Where to set proficiency levels and passing scores for each of the HSAs.
- Whether to expand the HSA program to 12 end-of-course exams, of which students would have to pass 10 to graduate
- How to establish student accountability for achievement without over-penalizing disadvantaged students

▲

*Is the increased
rigor behind
Maryland's diploma
worth the
probability that
fewer students will be
able to earn one?*

▼

- Ultimately, whether to make the HSAs a graduation requirement and, if so, when?

Currently, all high school students must take HSA tests in five courses, and normative test scores (how a student performs against others) are reported on their transcripts, yet proficiency levels measuring whether students have met course standards have not been set. As a result, neither the public nor State educators know whether high school students failed, met, or exceeded Maryland's expectations of what they must be able to do and know. One outcome, however, is clear: 2002 HSA data show predictable and significant achievement gaps on high-stakes tests between low-income minority students and their more affluent counterparts as well as between special education students and regular education students. Another outcome may be deduced: if Maryland enacts high-stakes testing as a graduation requirement, students at the lower end of the spectrum will be at risk for earning a diploma.

Meanwhile, the State has yet to articulate in any way how it will ensure, and

be held accountable for ensuring, that the HSA program amounts to more than a series of tests and truly improves opportunities for learning, achievement, and post-graduation success for Maryland's students.

**Difficult Decisions
with High Stakes**

It is up to the Board to ensure that the HSA program is conceived and implemented with the highest probability of producing improved learning and achievement for all students. The consequences of the Board's decisions for students are clear. If these are to result in all students receiving an adequate education, the Board's decisions require examination of the educational and moral dilemmas at the intersection of student, school and district accountability: how well will a system of sanctions and assistance improve education for all high school students, particularly if failing students are denied diplomas for not meeting standards? What impact will such high-stakes tests have on curriculum, instruction and learning in high school classrooms? What if the graduation requirement causes failing students to drop out, or causes schools to push out failing students, leading to higher HSA scores but higher dropout rates? Is the increased rigor behind Maryland's diploma worth the probability that fewer students will be able to earn one?

Finally, to ensure that accountability extends to the Board itself: how will the State measure and be held accountable for the degree to which the HSA program improves high school instruction and increases opportunities for learning, achievement, and post-graduation success?

Given the broad scope of the high-stakes testing debate and the myriad chal-

continued on page 3

continued from page 2

allenges encountered by participating states, there are no clear-cut solutions to the dilemmas Maryland’s education officials now face. This report aims to help the Board and its partners proceed thoughtfully and cautiously by presenting lessons learned about high stakes high school testing to date; and by raising difficult questions, including those that lead to scrutiny of the state’s own commitment to high standards, high-stakes, and meaningful accountability.

If the Board decides it is indeed committed to high-stakes testing, it must explore additional educational supports, second chances for test-takers, and alternatives to the HSAs for certain populations. It must also reconsider its test format and scoring procedures to ensure timely dissemination of test results and data as part of building a collaboration with districts, schools, students, and the public. An external evaluation tracking impact of the HSA program with continuous public feedback is critical.

The Evolution of Maryland’s HSA Program

High-stakes high school testing is not new to Maryland. In the late 1970s, the Functional Tests in reading, math, writing, and citizenship were designed to ensure that students acquired at least minimum competency in basic skill areas before graduating from high school. In 1983, the Board conditioned the award of high school diplomas on passing scores on these tests, beginning with the Class of 1987. By 1989, passing scores on all four Functional Tests were required for graduation, and remain so today for three of the four tests.²

But the Functional Tests are problematic. At a time when Maryland hopes to raise its high school standards,

the Functional Tests assess 6th-grade level knowledge and skills, and few data exist to show the value or impact of these tests. National studies suggest that minimum-competency exams have not improved academic rigor and standards, and Maryland’s experience appears to mirror those findings.

In 1989, the Commission on School Performance, appointed by then-Gov. William Donald Schaefer, noted the Functional Tests’ low standards in the influential Sondheim Report. This report recommended that Maryland implement a more demanding, comprehensive, and coordinated system of assessments, performance standards, and accountability, paving the way for programs that have since emerged in grades K-8, including the now-defunct MSPAP tests. It also recommended new criterion-referenced³ tests to measure academic standards within a

statewide accountability system.⁴ The HSA program, designed to replace the Functional Tests as a high school graduation requirement, was ushered in by the establishment of new high school standards known as Core Learning Goals in the mid-1990s.

Under the state’s original HSA plan, 12 end-of-course exams would accompany high school courses in the four major academic disciplines of English, social studies, science, and math. Students would have to pass 10 of the 12 in order to graduate. So far, exams exist for English I, Government, Biology, Algebra/Data Analysis, and Geometry, with students typically taking them in 9th or 10th grade as they complete the course. (The Geometry HSA, originally in Phase II, was adopted to meet the NCLB 10th grade math requirement.)

Maryland’s High School Assessment Program

| Subject Area | Course Credits Required | HSA Tests in Place | HSA Tests Projected to be Required for Diploma |
|---------------------|--|--|--|
| English | 4 | English I | English I English II English III |
| Mathematics | 3 Algebra I Geometry Other | Algebra/Data Analysis Geometry (Used for NCLB 10th grade math test) | Algebra/Data Analysis Geometry |
| Science | 3 Including one lab experience | Biology | Take two of the following: Earth/Space Science Biology Chemistry Physics |
| Social Studies | 3 Government World History US History | Government | Government World History US History |

continued on page 4

continued from page 3

The tests have a mixed-item format of “selected response” (multiple choice) items and “extended constructed response” items (requiring a written answer).

According to the Maryland State Department of Education, the more multiple-choice items, the more efficient the grading process and the faster the results, by a margin of at least four to six weeks. Yet, concerned that a purely “objective” test would force rote learning versus higher-level thinking skills, the State opted to include constructed response questions and to employ external graders, two decisions that severely limit the utility of these tests for school educators. “End-of-course”

exams are administered five weeks before the end of the school year, and, because scores are not reported until August at best, they cannot be used for school-based decisions about course grades, summer school eligibility, or promotion.

The HSA tests are offered in May and in January to accommodate varying academic schedules. Because the May administration of the “end-of-course” exam is five weeks before the actual end of school, districts often administer additional cumulative tests. Because the HSAs are externally scored and not entirely objective, the State expects to release May 2003 scores in August. This represents a shorter timeline than last year, when May 2002 results were released in December.

Even if the earlier date is met, it will not allow schools to use the tests for course grade promotion decisions, or permit students to retest or get instructional help during the summer.

Finally, students enrolled in HSA courses are required only to take, not pass, the tests to graduate. While percentile-ranked scores are reported on student transcripts (with the exception of Geometry), passing the HSAs is not currently a graduation requirement.

High-Stakes Testing: A National Challenge

While states have distinct policies driving high school reform, the common trend is toward greater student, school, and district accountability: Eighteen states now require students to pass exit exams prior to graduation; in six years, such tests are expected to exist in some 24 states.⁵

States moving toward greater accountability also face challenges that are forcing them to rethink their strategies. Maryland is not alone in delaying its proposed graduation requirement, nor should it be admonished for having done so, given the potential impact it will have on many students. Below is a sampling of states at similar impasses in implementing their graduation requirements:

Arizona postponed implementation of its requirement four times.

California just postponed its requirement for two more years.

Florida lowered its passing scores for math and reading, and waived the requirement for special education students.

Massachusetts lowered its passing scores and decided to issue waivers at the eleventh hour.

Samples of Question Types from the Algebra/Data Analysis Assessment Selected Response Question

1. Mary graphed the system of equations below.

$$y = \frac{3}{2}x + \frac{7}{2}$$

$$y = -\frac{2}{3}x + \frac{7}{3}$$

Which of these best describes the relationship between the two lines?

- A. They have no point in common
- B. They have one point in common
- C. They have two points in common
- D. They have infinite points in common

Extended Response Question

2. The table below shows the sales for a greeting card company for the years 1990 through 1998

| Year | Sales (in thousands) |
|------|-------------------------|
| 1990 | \$205 |
| 1991 | \$230 |
| 1992 | \$245 |
| 1993 | \$270 |
| 1994 | \$295 |
| 1995 | \$320 |
| 1996 | \$340 |
| 1997 | \$350 |
| 1998 | \$365 |

Complete the following in the Answer Book:

Write an equation for the line of best fit for this data. Let x represent the years since 1900 and y represent the sales, in thousands of dollars. (If you choose to draw a graph, use the grid provided in the Answer Book.)

According to your equation, what were the sales in 1999? Use mathematics to explain how you determined your answer. Use words, symbols, or both in your explanation.

continued on page 5

continued from page 4

Michigan created multiple diploma types to recognize different levels of achievement.

New York added a waiver/appeal process.

Wisconsin made its “requirement” optional for school districts.

Disproportionate Odds of Success: Graduation and Dropout Research

Research suggests a link between high-stakes tests, failure rates, and dropout rates, particularly for at-risk students. During the early years of high-stakes minimum competency tests in the 1970s, 20% of African-American students in Florida failed and were denied high school diplomas, compared to 2% of white students. More recently, researchers have asserted that high-stakes testing not only holds back low-income and minority students, but also is linked to dropout rates among them. Haney’s (2000) study of the Texas Assessment of Academic Skills found that dropout rates, namely among African-American and Hispanic students, increased as a result of the high-stakes exit exam in Texas.

In Massachusetts, meanwhile, where activists decry the negative effects of the Massachusetts Comprehensive Assessment System (MCAS), education officials assert that a declining dropout rate between junior and senior year (from 3.9% to 3.6%) proves the graduation requirement is helping students stay in school.⁶ Critics, however, contend the more likely factor is pressure to retain underperforming students prior to 12th grade before they drive up schools’ failure rates, or drive down measures of progress required for state and federal accountability. In Boston, 16% of the Class of 2002, which was not subject to the MCAS

exam, repeated 9th grade, compared to more than 25% of the Class of 2004, which must pass the MCAS to graduate.⁷

Student Accountability

On the other hand, high-stakes can spur student performance as well. A new study by Carnoy and Loeb (2003) shows no negative effect of student accountability on student retention or high school graduation rates.⁸ Focusing on 8th-grade achievement in math, it suggests that the higher the stakes in a state, the better the academic performance of its minority students. “If I were gambling on whether to put in a high-stakes system or not, I would put one in,” asserts study coauthor Martin Carnoy of Stanford University.⁹

Setting Cut Scores: A Volatile Mix of Science and Politics

One of the most critical and politically charged decisions state educators make in designing high-stakes tests is where to set the initial passing or “cut” score. Unless standards are set fairly low, there will likely be many students who do not meet them, spurring the question: Should standards be set low so fewer students fail? Or should they be set high to increase academic rigor, with the likely result that a politically unacceptable number of students will fail?

The process New York adopted illustrates the dilemma. It initially set cut scores low, at 55%, to avoid massive failures at disadvantaged schools, but then raised it to 65% over two to four years. As a result, more students reached the lower initial pass score in the first year, but fewer passed under the longer-term goal. In the first year, with the lower cut score, 72% of students in schools with largely minority populations passed; but only 44% of students in these schools passed using the higher long-term cut score.

Other states have also phased in standards and/or revised their cut scores over

time. Virginia adopted cut scores in fall 1998. After the first administration of its test, only 2% of schools met accreditation requirements; after the second administration, only 7% met standards. Public alarm ensued, and state education officials made changes in the quality and quantity of test content. Acknowledging the difficult standards of its own exit exam, Massachusetts adjusted its cut score downward; now students have to answer only 40% of the questions correctly to pass.

The Narrowed Curriculum

Critics of statewide testing in New York and elsewhere have charged that it has a negative effect on what is taught and how it is taught. A new study by J. J. Pedulla et al. (2003) finds that state testing programs have caused a “narrowing of the curriculum,” particularly in states with high-stakes for students and schools.¹⁰ Eighty percent of teachers in these high-stakes states report that pressure to produce high scores means that they teach little content beyond what will be tested. Such unintended consequences in the classroom, where a large-scale policy like the HSA program must ultimately work, are often the most difficult to identify.

Providing Alternative Pathways to Diplomas

Requiring passing scores on end-of-course exams for graduation raises questions about how failing students will relearn and retest. Massachusetts launched its graduation requirement for the Class of 2003 with no alternative options in place. It requires only two standards-based exams (English and math) and gives students five opportunities to pass them, starting in 10th grade. As of May 2003, when 8% of 12th-graders had not passed both tests a month prior to graduation, Massachusetts hastily decided to offer waivers.

continued on page 6

Meanwhile, many states have ensured that students who do not pass high-stakes tests have alternate means of earning diplomas:

- **Provide multiple options to retake the test** several times annually during and after leaving high school. In many cases, students are permitted to simultaneously pursue a higher-level course in the designated discipline. Some states with required graduation exams, such as Massachusetts, allow students to retest anywhere from two to 11 times, some up to the age of 21.¹¹
- **Offer a lower-level diploma** such as a Certificate of Completion. Alabama, Georgia, New York, New Mexico, North Carolina, and Vermont do this now.
- **Offer a proficiency assessment** for those who have failed high-stakes exams. Arizona has plans for a comparable equivalent demonstration test titled Arizona's Instrument to Measure Standards.
- **Establish a waiver or appeals process.** Students and their parents can make a case for graduating based on grade point average or other specified criteria in Indiana, Minnesota, and Mississippi. Massachusetts has granted more than 1,000 waivers to 12th-graders in the first year of its test requirement.
- **Provide multiple accommodation options for students with disabilities and English language learners.** A majority of states with testing programs, including Maryland, do this.

- **Substitute an alternative curriculum in lieu of tests.** Indiana's Core 40 curriculum is a rigorous high school-level curriculum that can be substituted for traditional graduation requirements for students who pass each course with a "C" or better.
- **Allow substitute test scores.** Some states accept other test scores, though this alternative is typically extended to high-achieving students. In New York and Virginia, students can substitute Advanced Placement, SAT II, and International Baccalaureate scores for the relevant content area's exit exam, thus relieving concerns about tests that "dumb down the curriculum."

Maryland's HSA Challenges

Other states' experiences may be useful to Maryland, but with high-stakes testing still in its infancy, there is little evidence of its effectiveness in improving educational opportunities. Maryland faces its own particular challenges, which the Board must weigh in moving forward.

External Challenges: Budgetary and Legislative

Two external factors influencing the course of Maryland's HSA program are the current uncertain economic times and the federal No Child Left Behind Act of 2001.

The early 1990s fiscal environment during which the HSA program was conceived held brighter prospects for education spending than the budget constraints of today. Uncertain state funding has threatened the Board's ability to finance necessary academic interventions for failing students, causing it to delay the graduation requirement three times, such that the requirement intended for students entering 9th grade in Fall 2001 is now proposed for those entering 9th grade in Fall 2004. The Board must consider the feasible scope of its program as budget

woes continue.

The requirements imposed by NCLB have also created challenges for the HSA program. Signed into law by President Bush in January 2002, NCLB aims to raise achievement among all students, and it gives states flexibility in setting standards and designing tests that measure those standards. But NCLB is less flexible about its goal of holding schools and districts accountable for ensuring that "no child"—or no group of children categorized by race, income, or ability—is left behind.¹² State officials warn that this lack of flexibility may require them to stretch limited state funds very thinly across huge numbers of schools designated as "needing improvement."

NCLB requires states to set reading, math, and science standards that all students must meet by the end of the 2013-14 school year. Schools must also have "average yearly progress" goals and bring each group of students—grouped by race and ethnicity, income, disabilities, and English proficiency—up to those levels. NCLB also requires student groups within schools to meet specified graduation rates. Schools missing annual or graduation rate targets over time for any subgroup will be labeled as "needing improvement" and required to target funds to low-performing students. This strikes at the heart of the Board's high-stakes dilemma, because research and HSA data show that more rigorous assessments affect graduation rates among disadvantaged and disabled youth in particular. If the Board opts for a graduation requirement and sets a high bar for passing, Maryland schools could disproportionately be identified as "needing improvement."

Internal Challenges: Limitations of the HSA Program Thus Far

No Indication of How Maryland Students Perform vs. Standards

continued from page 6

For students now taking the HSAs, scores are reported on their transcripts as a percentile ranking, meaning their performance on the test is compared with that of all other Maryland students taking the same test, not the percentage of questions the student answered correctly. Students have yet to be measured against objective standards—the purported basis of these criterion assessments. As a result, students, parents, teachers, and schools have no idea to what degree students failed, met, or exceeded state standards in HSA-required courses. Even though such information is slated to be available later this year, its absence now severely limits the opportunity to learn anything from HSA implementation.

Reporting Lag Limits Usefulness of Data

A report by Achieve, Inc. asserts that statewide testing results for high-stakes assessments must be given timely and informative release in order to be fully useful to students, parents, and educators. Maryland’s HSA program now falls far short of that goal.

Student data promised to districts and schools from the 2001 HSA field tests never arrived. The release of Spring 2002 data, slated for September 2002, did not occur until December 2002. Furthermore, State officials were unaware that Baltimore City Public School System, for one, had not released scores as of June 2003 to schools and students from the Spring 2002 HSAs, given a whole year earlier. This breakdown at the school district level signals a critical gap in execution at the State level. It appears the expected lag time between HSA tests and scores will continue to be about three to four

months, far too long for the results to be integrated meaningfully, if at all, with course grades, promotion decisions, and summer remediation.

Achievement Gaps Are Already Evident

The most recent HSA percentile rank scores from 2002 indicate performance gap patterns among students, schools, and districts based on race, family income, and special education status. Given the potential link between high-stakes tests and dropout rates, particularly among at-risk students, this is a critical gap for Maryland

education officials to understand.

Median percentile scores were reported for all Maryland schools and districts, and disaggregated by racial and ethnic groups, students with disabilities, English language learners, and low-income students who qualify for free and reduced lunch and Title I funds. The median is the middle score in a set of ranked scores, so that a group with a median percentile score above 50 is above average compared to all other Maryland students.

Despite existing testing accommoda-

Available 2002 HSA test data show the following:

Regular education students performed, on average, at the 53rd percentile, while students receiving special education services performed at the 19th percentile.

2002 HSA Performance

Regular education students and Students receiving special education services

| HSA score by median percentile | Algebra | English | Biology | Government | Geometry | Overall mean % |
|--|---------|---------|---------|------------|----------|----------------|
| Regular education students | 52 | 54 | 53 | 53 | 52 | 52.8 |
| Students receiving special ed services | 20 | 16 | 19 | 18 | 22 | 19.0 |

• • •

Low-income students performed, on average, at the 31st percentile, while more affluent students performed at the 54th percentile.

• • •

Asian-American and white students noticeably outperformed Hispanic-, African-, and Native American students, with the greatest disparity between Asian-Americans (73rd percentile) and African Americans (32nd percentile).

2002 HSA Performance by ethnicity

| HSA score by median percentile | Algebra | English | Biology | Government | Geometry | Overall Mean% |
|--------------------------------|---------|---------|---------|------------|----------|---------------|
| Asian | 75 | 71 | 74 | 68 | 76 | 73 |
| White | 63 | 61 | 63 | 61 | 61 | 61 |
| Hispanic | 39 | 41 | 41 | 41 | 40 | 40 |
| American Indian | 35 | 36 | 37 | 36 | 45 | 38 |
| African American | 28 | 34 | 30 | 34 | 33 | 32 |

• • •

Students in Baltimore City and Prince George’s County, whose populations are predominantly African-American (and, in Baltimore City, poor and urban), performed below students in suburban and rural counties, regardless of their poverty level.

continued on page 8

continued from page 7

tions for students with learning disabilities,¹³ their relative performance on the HSAs is extremely low. In Virginia, special education students are allowed to earn a “modified standard diploma” without having to pass its exit exams. This idea has been floated in Maryland as well, but parents and advocates worry that a different diploma would consign special education students to a second-class education. Whether this differentiated diploma would affect a student’s employability is unclear.

Gaps among races are also stark. Long-standing poor performance trends among minority and low-income students within districts and schools, and for majority-minority districts like Prince George’s County and Baltimore City, appear to continue unabated under the HSA program.

HSA Program Lacking Student Remedial Supports and Alternatives

A student who fails an HSA is required by the State to participate in “remediation” before retaking the test, yet what constitutes remediation has yet to be elaborated and could constitute another barrier. Furthermore, no graduation alternatives have been proposed to accommodate students who fail an HSA multiple times. When the Board considers its policies for retesting, test preparation, intervention, and remediation for students who do poorly on the HSA tests, it must strive to ensure that students are provided with the necessary academic supports and alternative assessment options to help them succeed.

Missing Links with Higher Education

Not only is it impossible to know from HSA test scores whether Maryland students are meeting State standards, it is also unclear whether they are meeting the expectations of post-secondary institutions. Despite Maryland’s high college

▲

***The most recent
HSA percentile rank
scores from 2002
indicate performance
gap patterns among
students, schools,
and districts based
on race, family
income, and special
education status.***

▼

remediation rates, post-secondary institutions have been noticeably absent from the HSA debate, and little has been done to link high school performance on the new HSAs to admission and remediation decisions at Maryland’s colleges and universities.

A Critical Lesson in Public Relations

Data measuring student performance against standards are slated for release soon, but the State appears reluctant to fully publicize student performance. Much of the MSPAP testing process was cloaked in secrecy, and Maryland reports its high school graduation rate as higher than that shown for Maryland in national studies. Now, the HSA program appears short on information-sharing. Two years into the HSA testing regimen, students have yet to receive a meaningful interpretation of their scores as compared to standards. More critically, the Board itself is preparing to make key decisions about high-stakes testing without the benefit of knowing how Maryland students actually fared on the HSAs. This lack of public information has long plagued public edu-

cation reforms in Maryland, and it is up to the Board to correct that problem as it considers whether to implement its highest-stakes assessment system ever.

Past and Present Delays May Derail the Original HSA Program

State education officials concede that they underestimated how long it would take to construct, validate, and field-test four HSAs. Just this spring the Board decided for the third time to delay the HSA graduation requirement by a year. These delays, plus the fact that continued budget woes could cause more postponement, demand that the Board carefully consider its pace moving forward.

Given the controversies surrounding Maryland’s past efforts with standards assessments and the high-stakes troubles of other states, the current cautious approach may make sense. On the other hand, delays allow testing opponents to create organized opposition. During the 2003 legislative session, the coalition Marylanders Against High Stakes Testing rallied behind the failed House Bill HB-1166, which would have barred the Board from requiring any assessment(s) for graduation from a public high school. Should the Board decide to adopt a graduation requirement and/or expand the HSA program to 12 tests, can it do so in a thorough yet timely fashion?

Recommendations

- 1. Establish high standards immediately, but raise the stakes for students on the High School Assessments gradually over time.**

Evidence suggests a negative effect of high-stakes testing on graduation rates among students who are disadvantaged or at risk. Based on what already appear to be achievement gaps in HSA scores, a graduation requirement would likely have

continued on page 9

continued from page 8

a disproportionately negative impact on those students who most need help. At the same time, setting low standards to avoid a decline in graduation rates would undermine the program's goal to enhance academic rigor.

Meanwhile, State funding for academic intervention for failing students (and instructional support in HSA courses) will likely remain limited over the near term. Given the long-standing deficits in many students' school and home environments, the notion that even a fully funded intervention plan will somehow magically create equal opportunity for all is flawed. Such challenges notwithstanding, the Board could fulfill the HSA program's mission of increasing rigor in one of two ways.

A) *It could adopt a medium-stakes alternative to the graduation requirement that would require all school jurisdictions to incorporate HSA tests as final exams comprising a significant*

part (at least 30 percent) of a student's final grade in that course. This option would require Maryland to change its scoring process in order to facilitate faster reporting.

B) *Alternatively, it could set a composite score for performance on all the HSA tests, whereby a student's good performance on one HSA would offset poorer performance on another.*

2. Put adequate and feasible remediation and alternative paths to a high school diploma in place before HSA-linked graduation requirements go into effect.

While the Board worries about imposing high stakes on students without providing the means for them to succeed, it has not publicly revealed how well Maryland's students have performed against objective standards. In the face of known performance gaps, schools without data on their students' objective performance cannot plan remediation options for students who fail one or more HSAs. Like other states, Maryland should provide alternative pathways to a diploma for students who have failed HSAs repeatedly using alternative assessments, a waiver process, alternative curricula, and/or differentiated diplomas. It should also consider whether some students with disabilities could pursue the current Independence Mastery Assessment Program diploma. At the other end of the spectrum, the Board could combat concerns about a less rigorous curriculum by accepting, in lieu of the HSAs, passing scores on nationally recognized tests typically taken by high-achieving students, e.g., the appropriate SAT II, Advanced Placement or International Baccalaureate exams.

3. Ensure that the HSA test design and scoring process will be useful

in improving high school instruction. At the least, insist on faster turnaround of test results.

It is not clear that the Board knows the timing, policy, and cost implications of various testing alternatives. The HSAs should be an integral component of instructional decisions like grading and promotion, and should function as a true cumulative end-of-course test. Revising the test format and/or scoring process would allow tests to be taken at the true conclusion of the course, and scores could be made available within 10 days. To achieve this, the State could use only multiple-choice items on its HSAs or have them scored by the high schools; either of these changes would be more constructive and less costly than current practice.

The Board should make explicit decisions about the HSA test format (whether to adopt all multiple choice questions, for example) after publicly presenting and debating the cost and timing implications of various policy options. The same is true with in-school scoring of the HSAs accompanied by external audits, a system that appears to be successful in New York State.

4. Assign similar student accountability to the two NCLB tests given in 10th grade.

While student accountability is not a component of the federal NCLB requirements, research suggests that attaching stakes to tests used for school and district accountability increases the likelihood of improved school performance measures, particularly among at-risk youth. Maryland should consider a student accountability measure in all NCLB high school tests and end-of-course exams. The composite score alternative cited above stands a good chance of spurring student effort without decreasing graduation rates.

continued on page 10

▲
***This summer,
the Board has an
opportunity to re-engage
the public by making
the process of scoring
tests, scaling scores,
and setting performance
standards more
transparent to those
most affected by
these decisions.***
▼

continued from page 9

Moreover, the state could integrate the 10th grade NCLB reading test with the proposed 10th grade English II HSA requirement.

5. Make the HSA program more transparent, starting with release of data on how Maryland students performed against standards.

The State has not provided adequate or useful data about student HSA performance, so stakeholders are ill prepared for potentially bad news about their failure to meet state standards.

This summer, the Board has an opportunity to re-engage the public by making the process of scoring tests, scaling scores, and setting performance standards more transparent to those most affected by these decisions. We recommend a public release of students' criterion-referenced (that is, objective) scores on the 2002 (and 2003, if possible) HSAs as a critical step in establishing cut scores,



Finally such a study would address what, ultimately, are the most important questions: As a result of the HSA program, will all students have received a better education? Will their post-secondary prospects be enhanced?



even if doing so delays final adoption of passing scores. We encourage the State to use community forums to gather feedback and share information about the graduation requirement. Finally, we urge the Board to ensure that members of panels setting cut scores represent Maryland's diverse HSA stakeholders.

6. Hire an independent evaluator to measure and publish data annually on HSA progress toward clearly articulated educational outcomes and to identify unintended consequences.

The Board should institute a process for setting quantifiable objectives, measuring desired student outcomes, and evaluating the HSA program in terms of student, school, and district progress toward these outcomes. Most importantly, the study should identify unintended consequences of the HSA program (e.g., dropout rates) and should examine the tests' impact on the entire high school curriculum. A longitudinal study by an independent evaluator (without a pro- or anti-testing ax to grind) could track progress toward outcomes, allow mid-course corrections in tests and improved support for schools and students, determine the program's effect on the quality of high school instruction and student learning, and inform instructional techniques of teachers. Finally such a study would address what, ultimately, are the most important questions: As a result of the HSA program, will all students have received a better education? Will their post-secondary prospects be enhanced?

7. Align performance on the HSAs with college admissions and placement in Maryland's post-secondary institutions.

The Board should engage higher education institutions in a meaningful way to

ensure that HSA course and test content aligns with post-secondary requirements. A stronger, coordinated program could inspire willingness among the State's higher education institutions to reward strong HSA performance with preferences in admissions and financial assistance. This could create an incentive for students to perform well, and eliminate the barrier of placement testing, particularly at Maryland's community colleges.

8. Delay expansion of the HSA program and integrate HSA and NCLB testing requirements.

Maryland debuted five high-stakes exams at once, a major feat that was complicated by NCLB's mandate to implement reading and math assessments for all 10th grade students in 2003. We suggest that Phase II of the HSA program be delayed until findings from an external evaluation of Phase I (Recommendation 6) are communicated. We also suggest that HSA and NCLB testing requirements be integrated so student accountability is consistent across all school and district accountability measures.

Conclusion

Maryland is at a critical juncture as policymakers decide how high to set the bar for graduating from high school: whether to make passing scores on the HSAs a graduation requirement; what exactly constitutes passing the tests (i.e., what proficiency standards students must meet and how those translate into passing scores); and whether to proceed with plans to introduce up to seven more tests.

The Board can decide to move quickly to raise the standards and stakes for a diploma, which will likely result in thousands of 12th grade students being unable to graduate in May 2008, particularly those who are minorities, impoverished,

continued on page 11

▲

Will the High School Assessment program as currently envisioned achieve the State's goal of producing high school graduates with levels of skill and knowledge that greatly enhance the value of Maryland's diploma?

▼

disabled, or otherwise at risk or disadvantaged. Or it can choose a more judicious path of relying on high standards with gradually increasing student consequences, a more efficient test format and scoring procedure, an external evaluation with annual reporting, mid-course corrections, and greater public and higher education engagement to improve the quality of high school learning across the board.

Will the High School Assessment program as currently envisioned achieve the State's goal of producing high school graduates with levels of skill and knowledge that greatly enhance the value of Maryland's diploma? The Board must answer this high-stakes issue in an evolving and uncertain environment by choosing the path of what is best for Maryland's students. ■

References

- Amrein, A., and Berliner, D. (2002). High-Stakes Testing, Uncertainty, and Student Learning. *Education Policy Analysis Archives*. 10(18). Retrieved from <http://epaa.asu.edu/epaa/v10n18/> .
- Amrein, A., and Berliner, D. (2002). *An Analysis of Some Unintended and Negative Consequences of High-Stakes Testing*. Tempe: Arizona State University.
- Bishop, J. (1998). The effect of curriculum-based external exit systems on student achievement. *Journal of Economic Education*, 29(2), 171-82.
- Carnoy, M., and Loeb, S. (2002). Does External Accountability Affect Student Outcomes? A Cross- State Analysis. *Educational Evaluation and Policy Analysis*, 24(4), 305-321.
- Chadowsky, N., Kober, N., Gayler, K., and Hamilton, M. (2002). *State High School Exit Exams: A Baseline Report*. Washington, D.C.: Center on Education Policy. Retrieved from <http://www.ctredpol.org/pubs/statehighschoolsexitexams.html>
- Clarke, M., Haney, W., Madaus, G. (2000). High Stakes Testing and High School Completion. *NBETPP Statements*, 1(3), Chestnut Hill: Boston College, Center for the Study of Testing.
- Greene, J. (2001; April 2002 Revision). *High School Graduation Rates in the United States*. New York: The Manhattan Institute for Policy Studies.
- Hamilton, L., Stecher, B., and Klein, Stephen P (eds.). (2002). *Making Sense of Test-Based Accountability in Education*. Santa Monica: Rand Publications.
- Haney, W. (2000). The Myth of the Texas Miracle. *Educational Policy Analysis Archives*, 8(41). Retrieved from <http://epaa.asu.edu/epaa/v8n41/> .
- Horn, C. (2003). High Stakes Testing and Students: Stopping or Perpetuating a Cycle of Failure? *Theory into Practice*, 42(1), 30-41.
- Jacob, B.A. (2001). Getting Tough? The Impact of High School Graduation Exams. *Educational Evaluation and Policy Analysis*, 23 (2), 99-121.
- Kellaghan, T., Madaus, G., and Raczek, A. (1996). *The Use of External Examinations to Improve Motivation*. Washington, DC: American Educational Research Assoc.
- Le, V. (2002). *Alignment Among Secondary and Post-Secondary Assessment in Five Case Study States*. Santa Monica: Rand Publications. Retrieved from <http://www.rand.org/publications/MR/MR1530.0/> .
- Linn, R.L. (2000). Assessments and Accountability. *Educational Researcher*, 29(2), 4-16.
- Maryland State Department of Education. (2002). *Aiming Higher: The Next Decade of Education Reform in Maryland*. Washington, D.C.: Achieve, Inc. Retrieved from http://www.msde.state.md.us/pdf_files/aiming.pdf .
- McDermott, K. (2001). *State Testing and Graduation Requirements: A Framework for Analysis*. Amherst: Center for Education Policy, University of Massachusetts.
- Raymond, M., and Hanushek, E. (2003). High Stakes Research. *Education Next*, (3) 48-54.
- U.S. Office of Technology Assessment (1992). *Testing in American Schools: Asking the Right Questions*. Washington D.C.: U.S. Government Printing Office.

Footnotes

- ¹ Maryland State Department of Education, *Maryland High School Assessments*, October 2001, <http://www.msde.state.md.us/hsimprovement/pdf/HSAPPOct2001v1.ppt>
- ² As of the 1998-99 school year, the Citizenship test was deleted as a graduation requirement
- ³ Criterion-referenced tests are tests that measure student achievement against a specified body of knowledge. They are typically contrasted with norm-referenced tests, which measure student performance in comparison with the other students who took the test.
- ⁴ *The Report of the Governor's Commission on School Performance*, August 1989, p. 8-10, 15-17.
- ⁵ CEP, 2002. Note that CEP excludes states that have tests but where those tests are not required statewide, that are differentiated diploma systems, or where student accountability is not a part of the system.
- ⁶ Kevin Rothstein, "MCAS credited for state's decreasing dropout rate." *The Boston Herald* (February 14, 2003). BostonHerald.com
- ⁷ Rothstein.
- ⁸ Raymond, M. and Hanushek, E. (2003). High Stakes Research. *Education Next*, Summer 2003 www.educationnext.org.
- ⁹ Winter, G. "New Ammunition for Supporters of Do-or-Die Exams." *The New York Times* (April 23, 2003).
- ¹⁰ Pedulla, J., Abrams, L., Madaus, G.F., Russell, M.K., Ramos, M.A. and Maiao, J. (2003). "Perceived Effects of State-Mandated Testing Programs on Teaching and Learning: Findings from a National Survey of Teachers." Chestnut Hill, Mass: National Board on Educational Testing and Public Policy, Boston College.
- ¹¹ CEP, p.48
- ¹² Neill, M. "High Stakes, High Risk," *American School Board Journal* (February 2003).
- ¹³ "Accommodations will be consistent with those provided during classroom instruction and testing and designated in the IEP for a student with a disability, or in the plans for eligible 504 or limited English proficient students." Message to Maryland Parents from Dr. Nancy Grasmick, December 2002. Retrieved from <http://www.msde.state.md.us/testing/Dec2002HSAletter.pdf>

High Risk or High Time?
is available on The Abell Foundation's
website at www.abell.org

ABELL SALUTES:

Continued from page 1

tem to measure results of the intervention. Eighty repeat victims who agreed to participate were randomly placed into one of two groups, the intervention group where 40 participants received a comprehensive range of support services, including drug treatment, counseling, job training and placement, and the control group, where participants received only the minimal services provided to repeat victims.

Data gathered after three years revealed significantly positive results for the intervention program: the control group was three times more likely to be re-hospitalized than the intervention group, and was more than two times more likely to be convicted of a violent crime during the period of the intervention, with a high percentage of individuals being sent back to prison. Only 20 percent of the control group were employed, as opposed to 82 percent of the intervention group.

A strong evaluation component focused on cost-effectiveness of the intervention, and a three-month, six-month and one-year follow-up of the participants in both groups. The hypothesis being tested was that a cost effective hospital-based intervention program reduces violent recidivism and crime, and leads to significant economies. Using the \$42,000 per admission, in any given year, as 10 re-admissions are eliminated, the savings to supporting agencies, including Maryland taxpayers, is \$420,000.

But for the victims in VIP, the program is more about turning life around: One writes, "I had been shot, my life had no meaning. I was introduced to VIP—and my whole life changed." And another, from prison, "I will be putting my full attention to the VIP program just as soon as I am released. It's nice to know you're in my corner." And, "I came into Shock Trauma from a gunshot wound. Since, VIP helped me with counseling, and drug addiction program. VIP is a shot of hope for innercity youth."

The Abell Foundation Salutes VIP, its Program Coordinator, Dawn Esslinger, and the entire staff, for operating a program that saves victims' lives and taxpayers' money. ■