

(De)Grading the Standardized Test

Can Standardized Testing Evaluate Schools?

NICOLA E. SIMMONS

In late November 2003, the *Burlington Post* reported, “Halton’s public schools holding their own in provincial testing”.¹ Was there much rejoicing among regional schoolteachers that day? Did pressure ease, did parents smile, did principals congratulate; in short, was it viewed as a great achievement of teaching success? And if so, is it the whole picture? Do high marks on large-scale standardized tests prove that students are doing better in school? Do they indicate which are the better schools?

“Accountability!” runs the education battle cry. “Test the troops!” Is the reply.

Standardized testing is an assessment strategy that evaluates all students and all schools on the same basic skills and, therefore, might reasonably indicate which schools are high performers and which are not. Standardized testing is not a new strategy for providing this proof. The United States has produced reams of articles criticizing standardized testing and its effect on students, teachers, and education. Other countries, Canada being one, are jumping on the standardized testing bandwagon, leaving many wondering whether due consideration has been given to the American experience.

This article looks at standardized testing in the United States and asks these questions: Do standardized tests measure curriculum outcomes? Do they accurately reflect student learning? And finally, do standardized tests really tell us which are the best schools?

(De)Grading Outcomes

Why do we give tests? Evaluations tell students whether they have achieved the course outcomes, providing them with positive rewards for success, encouraging them to do more, and indicating areas needing more work. Testing also provides feedback about teaching: whether enough emphasis has been given to major concepts and



whether explanations have made sense to the students. Standardized tests are designed to establish an educational guarantee that, before moving to the next level, students have achieved the outcomes necessary for success. Additionally, these tests can provide information about school performance and where education dollars could be spent to develop programmes.

Standardized testing at its best could determine whether each student has met the desired outcomes and is ready to go on to the next level. Just as a driving test ensures that everyone who drives knows the rules of the road and can drive safely, standardized testing could make sure that students have mastered the basics. A number of professionals, such as doctors, lawyers, accountants, real estate agents, and others must write standardized tests before being allowed to enter practice. Such tests can be designed to include a core set of objectives on which the student must score 100%, because a 60% in surgery just doesn’t cut it, so to speak.

While this sounds reasonable, the problem with current standardized tests is that the standards being tested are often separate from those reflected in the curriculum. Even when there are matches, the outcomes assessed by standardized tests are seldom in harmony with the outcomes intended by the curriculum. In fact, one study from the University of Wisconsin found that in one subject the overlap was just 5 percent.²

James Popham, a retired standardized-test creator, tells how this problem arises. Test writers, he says, make positive claims about the reliability of their tests to make it more likely that the test will be purchased. The higher the score spread, the easier it is to prove reliability. This creates the problem: The test writer’s aim is to create a spread of marks by asking questions not every student can answer. As a result, “standardized achievement tests should not be used to evaluate the quality of students’ schooling because the quest for wide score-spread tends to eliminate items covering important content that teachers

In opposition to the cognitive multi-level approach of the classroom activities, standardized tests focus on the final step rather than the learning path, and in doing so, may ignore higher-level cognitive skills.

have emphasized and students have mastered”.³ In other words, questions covering content that has been a focus of classroom teaching, and that students are likely to get correct, are removed from the test. The result is testing that does not accurately reflect the prescribed curriculum.

Eventually, of course, there will be a match: as greater emphasis is placed on accountability and testing, teachers will work even harder to guarantee student success – on those tests. One teacher speaks of the challenge inherent in preparing students for testing. “When school began in the fall, we had six weeks to prepare our sophomores for the test. Teachers were encouraged to abandon the regular curriculum and instead devote time to test practice.”⁴ In short, rather than planning assessments that evaluate the curriculum, teachers are planning curriculum to match the assessments. The outcomes achieved are no longer those specified by the approved curriculum, and students are being taught how to take tests rather than how to learn.

(De)Grading Student Learning

As teachers are pressured to improve test scores, and as tests increasingly become the classroom focus, what type of learning is being emphasized in our schools? Certainly students are learning new skills through standardized testing. A writing teacher coaches her students towards test success. “Don’t be too creative. Don’t think too hard. Only give them what they want. Don’t risk. Pace yourself for limited space and time.”⁵ It’s easy to anticipate the gaps this will leave in students’ education. Critical thinking and creative problem solving, those essential skills in our modern world, will be less and less valued. Another teacher bemoans the results. “We’ll have very successful third-grade readers...who can pass tests...but they also need to think creatively, communicate effectively and work collaboratively.”⁶

Today’s curriculum asks students to perform a myriad of tasks at different cognitive levels. Classroom activities engage students in higher order tasks of analysis, synthesis and evaluation. Standardized testing does just the opposite. “Children who have been taught to construct their own answers to problems...are confronted

EN BREF Les tests standardisés évaluent chez les élèves de toutes les écoles les mêmes aptitudes fondamentales et, par conséquent, il peut sembler juste de les utiliser pour effectuer des comparaisons entre les écoles. Mais, ces tests peuvent-ils mesurer les résultats du programme d’études et reflètent-ils avec précision l’apprentissage de l’élève? Nous permettent-ils réellement de déterminer quelles sont les meilleures écoles? Or, les tests standardisés ne seront utiles et valables que lorsqu’ils nous permettront d’évaluer ce qui a été enseigné dans la salle de classe, qu’ils poseront des questions neutres qui portent sur ce que les élèves ont appris, qu’ils encourageront le recours à des démarches individuelles et distinctes pour trouver les réponses et qu’ils serviront à améliorer le programme d’études plutôt qu’à stimuler la concurrence entre les écoles.



with someone else’s answers and are told that there is only one correct answer.”⁷ In opposition to the cognitive multi-level approach of the classroom activities, standardized tests focus on the final step rather than the learning path, and in doing so, may ignore higher-level cognitive skills.

There is always a concern that any single assessment risks an inaccurate view of the students’ knowledge and skills. On any given day, some students will be sick, overly tired, depressed, or simply unhappy. As any teacher can confirm, students bring personal issues with them into the classroom. In this respect, they are no different from the rest of the population. “The stress induced in students by the conditions of test-taking and the consequences for test ‘failure’... can become self-perpetuating test anxiety, which can cause students to ‘freeze’ or ‘go blank’ and forget information they know.”⁸ Some of these difficulties are acknowledged in a statement from the American Education Research Association (AERA), which cautions against “relying solely on test scores to determine whether a student should graduate or move on to the next level.”⁹ However, standardized test scores seem to be used to do just that. In Ontario, grade 10 students must pass a literacy test or they cannot graduate from high school.

If better test scores require spending more time in test preparation, teachers will respond to that need, even though it may come at the expense of other learning. Nor does this escape student notice. One student, referring to the statewide test given in South Carolina, comments, “All they care about is the...test; they don’t care if we learn anything.”¹⁰ Success at taking tests becomes more important than success at learning.

(De)Grading Schools

Standardized testing results are currently being used in real estate adver-

No longer is a house desirable for its proximity to a school; it must be a good school as proven by standardized tests.

Accountability should be applauded: accountability to students, to teachers, to schools, to society, to the entire process of education. However, it is not clear with standardized testing: exactly who is accountable for what?

tising. No longer is a house desirable for its proximity to a school; it must be a good school as proven by standardized tests. Do the results of these tests really tell us which are the best schools?

In fact, rather than assessing student and school performance, standardized testing does little more than confirm the socioeconomic status of students being tested, and in doing so, it discriminates against certain schools. Alfie Kohn, a critic of standardized testing, writes that such tests are “unfair because the questions require a set of knowledge and skills more likely to be possessed by children from a privileged background. It’s more than a little ironic to rely on biased tests to address educational inequities.”¹¹

The reward system for high-scoring schools may be a strong contributing factor in the push for standardized testing, and budget may not be all that is affected. Are strategies being used to exclude students who wouldn’t score well from writing the test? “Performance of principals and superintendents is often linked to the number of students who reach the bar set by the state. Unfortunately, this may mean that it is in an administrator’s best interest to eliminate systematically from the test those students who probably wouldn’t make the bar... some schools have relied on increased special education placement and retention in grade.”¹²

The craving for accountability has not resulted in a testing system that provides useful data about schools or their success. In New York, a study by the state found that “high school marks were better predictors of university marks than commercially or bureaucratically prepared tests similar to those for which there appears to be such a clamour.”¹³

At this point, one might wonder why high-stakes testing is used at all. The desire for accountability is at the root of the trend. As the 1998 Council of Ministers of Education (Canada) report says, “Accountability is a key priority.”¹⁴ Accountability should be applauded: accountability to students, to teachers, to schools, to society, to the entire process of education. However, it is not clear with standardized testing: exactly who is accountable for what?

“Massachusetts has spent \$20 million on the development of the Massachusetts Comprehensive Assistance System, which will be required of all tenth-graders beginning in 2003.”¹⁵ There is no question that money can be spent to improve education; however, one might well question the value of purchasing a \$20 million dollar rubber stamp that says “Not good enough”. In short, students may have become pawns in an educational game to grab the available glory and funding dollars. In Canada, “Alberta’s version of outcomes-based education...links cash bonuses in return for improved test scores at a time of budget cutbacks.”¹⁶ Perhaps this money would be better spent enhancing programmes at schools that have not scored so well, in order to truly improve student learning.

(De)Grading Assessment

Just as students, teachers, and schools need to be assessed, the tests themselves need to be assessed. We must not only test curriculum outcomes, we must create outcomes for the tests themselves, in order to verify that the process is worthwhile. A former president of the Carnegie Foundation for the Advancement of Teaching is quoted as saying, “The key to problem solving is not to figure out the answer to a question that someone else hands you, but to define the right problem. An educated person today is someone who knows the right questions to ask.”¹⁷ Let us ask these questions about standardized testing:

1. Does the test match the enacted curriculum, such that the questions on the test evaluate what has been stressed in the classroom?
2. Does the test pose unbiased questions that focus on what students have learned, and does it encourage individual and diverse approaches to answers?
3. Are test results used for programme enhancement rather than school competition?

When the above questions can be answered with an affirmative ‘yes!’, only then will standardized testing receive a passing grade. ★



NICOLA E. SIMMONS is a graduate student and B.Ed. facilitator in the Faculty of Education at Brock University and a part-time professor at Humber College in the Professional Development Department. Her education interests are wide-ranging, including particular interests in learning styles, participant-centred learning, and individualized assessment strategies.nicola.simmons@humber.ca

Notes

- 1 Tim Whitnell, “Halton’s Public Schools Holding Their Own in Provincial Testing,” *Burlington Post*, 28 November 2003, p. 4.
- 2 Ulrich Boser, “Teaching to the Test?,” *Education Week*. (2001): 1. [Retrieved 16 September 2002, from www.edweek.com/ew/ew_printstory.cfm?slug=39align.h19.]
- 3 James W. Popham, *The Truth About Testing: An Educator’s Call to Action*, (Alexandria, Virginia: Association for Supervision and Curriculum Development, 2001), 48.
- 4 Shannon Hughes and Jason Bailey, “What Students Think About High-stakes Testing,” *Educational Leadership* 59 (2002): 74.
- 5 Nancy M. McCracken and Hugh T. McCracken, “Teaching in the Time of Testing: What Have You Lost?” *English Journal* 91 (2001): 30.
- 6 Nan Youngerman, “When Vision, Not Test Scores, Sets School Priorities,” *Language Arts* 79 (2002): 228.
- 7 Diane E. Meaghan and François R. Casas “On the Testing of Standards and Standardized Achievement Testing: Panacea, Placebo, or Pandora’s Box?” *Interchange* 26 (1995): 48.
- 8 Robin H. McBee, “When it Comes to Testing, Why Not Make Lemonade?” *The Educational Forum* 66 (2002): 241.
- 9 Debra Viadero, “In Short” *Education Week*. (2000): 1. [Retrieved 16 September 2002, from www.edweek.com/ew/ew_printstory.]
- 10 P. L. Thomas, “Standards, Standards Everywhere, and Not a Spot to Think,” *English Journal* 91 (2001): 63
- 11 Alfie Kohn “Poor Teaching for Poor Kids”, *Language Arts* 79 (2002): 251.
- 12 Kathleen A. Steeves, Jessica Hodgson and Patricia Peterson, “Are We Measuring Student Success with High-stakes Testing?,” *The Educational Forum* 66 (2002), : 233.
- 13 Meaghan and Casas, 41.
- 14 Council of Ministers of Education, Canada, “Education Initiatives in Canada, 1998: A Report from the Provinces and Territories,” (1998): 27.
- 15 Steeves et al, 231.
- 16 François R. Casas and Diane E. Meaghan, “Renewing the Debate Over the Use of Standardized Testing in the Evaluation of Learning and Teaching,” *Interchange* 32 (2001): 148.
- 17 Steeves et al, 230.