

## **Teaching To the Test: What Every Educator and Policy-maker Should Know**

by [Louis Volante](#), Concordia University

### **Introduction**

The job of any teacher is first and foremost to promote learning in their students. Student learning should emphasize applied learning and thinking skills, not just declarative knowledge and basic skills (Jones, 2004). Ideally, students are able to develop the skills necessary to take what they have learned and apply this knowledge in a novel situation. In this sense, teachers are promoting authentic learning within their classrooms. In North America, however, high-stakes testing procedures have interfered with these goals, and are increasingly used to measure student knowledge and gauge the effectiveness of instruction. Each spring teachers throughout Canada are required to administer a series of provincially mandated tests to students in their classrooms (Simner, 2000). These standardized tests are often used to make comparisons across students, schools, and boards of education. Individual teachers and schools are often blamed for poor test results, which are typically reported in the press.

Standardized tests, when used appropriately, help teachers identify student strengths and weaknesses (McMillan, 2000). A standardized test is one that is administered and scored under uniform and controlled conditions (Payne, 2003). Used most commonly in K-12 schools, standardized tests are intended to measure learning outcomes and skills that are common to the curricula in a vast number of schools and school districts (Chatterji, 2003). Students typically complete norm-referenced tests that compare their performance to a representative sample of students in a norm group (e.g., a group of students at the national, regional, or provincial/state level) (Gronlund, 2003). Other students undertake criterion-referenced tests that compare their performance to a preset standard of acceptable performance in a particular area (Borich & Tombari, 2004). Both norm-referenced and criterion-referenced test results are increasingly used as benchmarks of success in North American schools.

Despite calls for the tracking of school performance with a broad range of outcomes, standardized tests continue to be touted as the most important measure of student performance (Mitchell, 1997). A recent study of student assessment in Canada called for a broad level of testing in all provinces in order to measure and improve school and student performance (Taylor & Tubianosa, 2001). Similarly, in a recent survey of U.S. public attitudes about education reform, almost one-half of the respondents noted the benefits of these tests for improving education (Hart & Teeter, 2001). Thus, despite the limitations of standardized testing, the majority of the public continues to view this type of assessment as an essential mechanism for measuring school performance.

Standardized test results often connect to high-stakes accountability for students, teachers, and schools (Linn, 2000). For example, students are required to pass the grade 10 literacy test in Ontario in order to receive a high school diploma. Assessment results in this province are also published in local newspapers ranking schools from highest to lowest on the basis of grades 3, 6, 9, and 10 results. Similarly, recent reforms precipitated by the No Child Left Behind Act (NCLB) in the United States have expanded testing and toughened standards for schools, teachers and students. Signed into law in 2002, NCLB requires all elementary students be tested annually from grades three through eight, with the public release of results. Schools that fail to meet Adequate Yearly Progress (AYP), as reflected in mandated improvements in test scores, are labeled as "failing". Although pressure on administrators, teachers, and students to meet high academic standards as reflected in high test scores can lead to productive work for many, the administration of standardized tests continues to raise important challenges for the teaching profession (Stiggins, 1999). Research suggests that teachers will often skew their efforts in the direction of activities that would lead to increases in these highly publicized scores (Earl, Levin, Leithwood, Fullan & Watson, 2003). Thus, teachers need to be especially weary of maladaptive preparation strategies that may accompany high-stakes testing measures.

Faced with increasing pressure from politicians, school district personnel, administrators, and the public, some teachers have begun to employ test preparation practices that are clearly not in the best interest of children. These activities may include relentless drilling on test content, eliminating important curricular content not covered by the test, and providing interminably long practice session that incorporates actual items from these high-stakes standardized tests (Popham, 2000). There have even been documented cases in the United States where teachers and administrators had given students the answers to standardized reading and mathematics questions (Goodnough, 1999). These cases of improper behavior in New York, Texas, Massachusetts, Maryland, Ohio and Connecticut were a direct result of increased pressure on schools that resulted from public rankings (Simner, 2000).

Test items are typically released so that teachers and students have a better understanding of the content covered by the test as well as the test's general format. For example, the Education Quality Accountability Office (EQAO) in the province of Ontario posts previous test items on their official website from their grades 3, 6, 9, and 10 assessments. Teachers are able to download copies of these previous assessments to help prepare for the upcoming testing session. Some teachers provide exercises featuring "clone items" – items so similar to the test's actual items that it's difficult to tell which is which (Popham, 2001). Focusing excessive classroom time on the teaching of released and cloned items can be construed as teaching to the test. This practice of teaching to the test has been noted in countries such as Canada, United States,

England, Australia, Japan, Israel and the Czech Republic (Levinson, 2000).

### Problems with Teaching To the Test

Few would debate the utility of providing teachers and students with information on the format and structure of standardized tests. Even knowledgeable students could miss an item (or a set of items) if they do not understand the mechanics of taking a particular test (Mehrens, 1989). It seems essential that teachers understand what is an appropriate amount of time needed for test familiarization, so that they do not sacrifice important curricular content in the drive for high scores. Clearly a week of mock exams prior to the administration of a test is excessive. However, even a day of practice testing is inappropriate if students are being trained how to answer particular test items. Popham (2001) has argued that item-teaching, instruction around items either found on a test or a set of look-alike items, is reprehensible since it erodes the inferences we can make about students' scores. Thus, appropriate test preparation is a question of both the amount of time and type of activities in which students are being asked to engage. An hour or two of class time that focuses on the structure and format of a test seems reasonable.

Teaching to the test also has a "dumbing" effect on teaching and learning as worksheets, drills, practice tests and similar rote practices consume greater amounts of classroom time (Sacks, 2000). Insofar as standardized tests assess only part of the curriculum, time spent on test taking often overemphasizes basic-skill subjects and neglects high-order thinking skills (Herman, 1992). Research suggests that while students' scores will rise when teachers teach closely to a test, learning often does not change (Shepard, 2000; Smith & Fey, 2000). In fact, the opposite may be true. That is, there are examples of schools from New York and Boston that have demonstrated improvements in student learning while their standardized test scores did not show substantial gains (Neil, 2003b). In both jurisdictions, these schools did not focus on teaching to the test or participating in test-preparation programs. Teachers that address the entire curriculum, particularly when preparing their students for standardized tests, provide their students with a solid foundation for future success.

Teaching to the test not only reduces the depth of instruction in specific subjects but it also narrows the curriculum so that non-tested disciplines receive less attention during the school day. Time is often devoted away from subjects like physical education, music, and drama so that teachers can provide more instructional time on commonly tested areas like reading, writing and arithmetic. Mary Lemon, a second grade teacher, succinctly summarized the pressures of testing and its effect on test preparation.

Everything that has to do with the test has been given such a high priority, that there is no priority any more but that ... The bottom line question comes down to, "Well, what's going to help them do better on the test?" And if it's not going to help them do better on the test, well, we don't have time for that right now (Wright, 2002, p.10).

Teaching a narrow curriculum is likely to alienate a large portion of students whose academic strengths lie outside of commonly tested subjects. It is important to note that this narrowing is likely to be greatest in schools serving at-risk and disadvantaged students, where there is the most pressure to improve test scores (Herman, 1992). One may argue that teaching to the test in such schools facilitates disengagement and even truancy.

Teaching to the test also undermines the validity of large-scale assessment results. Consider two schools: one which focuses test preparation on instruction with released or cloned test items; the other which provides instruction in a general body of knowledge that a test represents. Elevated scores by the former school likely do not represent authentic learning and students may not be able to fully utilize the skills that the test represents. This is not surprising, given that teaching to the test emphasizes memorization rather the application of new skills and knowledge in a novel situation. Thus, the predictive validity of a standardized test is compromised when teaching to the test techniques are employed (Burger & Krueger, 2003). Stripped of its power to make inferences about student skills and knowledge, a standardized test also loses its ability to inform instruction as a formative assessment measure.

Teaching to the test can lead to weaker and possibly incorrect interpretations about school programs (Mehrens, 1989). Smith and Fey (2000) note that practicing content known to be on the test can make a school look half a year better than a comparable school that did not employ teaching to the test preparation. Thus, schools may be mistakenly categorized as high achieving because of their utilization of inappropriate test preparation activities, not necessarily because of the actual characteristics of their student body. Funding may also be misallocated based on test results that do not include key aspects of learning (MacDonald, 2001). In such situations, the misdirection of funds may occur in school systems that are likely to be already struggling financially.

Another danger with teaching to the test is the negative consequence it has on the teaching profession as a whole. Stiggins (1999) astutely points out that the pressure to do well on high-stakes tests can sometimes have exactly the opposite effect from the one we see. Educators in North America rail against the amount of instructional time used to prepare for and administer tests (Levinson, 2000). Teaching to the test only serves to exacerbate feelings of frustration and disillusionment with the entire testing process. Weber (2002) documented the effects of high-stakes testing and the increased prevalence of teaching to the test in an inner-city California school. One teacher summarized her frustration with the schools test driven agenda by commenting.

The most pathetic thing is that up until two years ago, I counseled young people, "Come into teaching. It is a wonderful profession." Now I counsel them to find something else because this is not the profession I would choose for myself (Wright, 2002, p. 28).

Not surprisingly, the singular focus on standardized testing and the increased prevalence of teaching to the test resulted in an unhappy teaching staff that began questioning their suitability for the profession.

In addition to the previously cited problems to the breadth and depth of the curricula, teaching to the test may also erode basic skill development even in the tested subjects. Neil (2003a) reported cases where children have been taught to read by learning to look at the answer options to multiple-choice questions and then search the short passage to find the clue to selecting the correct answer. Independent evaluators have found that these children cannot explain what they have just read even though they got the test item correct. The implication is that there may be a significant number of test wise students who lack the basic skills needed to be successful in higher education settings.

Given the previous points, it seems logical that teaching to the test provides students with a skewed measure of their ability. Artificially high scores may lull students into a false sense of security, particularly for those heading to post-secondary institutions. If instruction focuses too heavily on the standardized test, students may not learn the skills they need to be successful in university such as making an oral presentation, conducting a science experiment, or writing a research report. Unfortunately, none of these skills are assessed in norm-referenced or criterion-referenced standardized tests. It is worth noting that there are schools that have avoided the temptation to teach to the test and been able to demonstrate lasting success for their students in terms enrollment and success in college (Neil, 2003b). Thus, individual teachers and schools can demonstrate their effectiveness without focusing heavily on standardized testing or resorting to teaching to the test techniques.

### **Alternatives To Teaching To the Test**

How can teachers avoid the temptation to spend excessive amounts of time preparing for standardized tests and focusing on teaching to the test techniques? The latter is no small task, especially when one considers the prospect of being judged against colleagues and schools that employ such methods. Undoubtedly, accountability for questionable test preparation strategies begins with individual teachers, schools and colleges of education. Both new and experienced teachers need to become cognizant of the dangers of teaching to the test and be instructed on constructive test preparation activities that promote authentic student learning.

Preservice classrooms and inservice workshops need to clearly outline unethical test preparation activities for aspiring and practicing teachers. Providing practice or instruction on a published parallel form of the same test or providing practice or instruction on the test itself is unethical (Mehrens & Kaminski, 1989). Teachers should receive training in curriculum-teaching which requires them to direct their instruction toward a specific body of content knowledge or a specific set of cognitive skills represented by a given test (Popham, 2001). Thus, teachers need to discuss test-represented content rather than test items themselves when preparing students for the challenges of high-stakes tests. Armed with better content knowledge and instructional strategies, teachers will be able to offer their students more than repetitive drills in preparation for taking standardized tests. These students will also be more likely to apply their new thinking skills and content knowledge in areas that extend beyond the confines of a particular test.

School administrators should also receive similar training if they are to be effective leaders within their schools. An administrator must be able to distinguish between appropriate and counterproductive test preparation strategies, and be able to convey this knowledge to their school staff. Through regular classroom visits, administrators should counsel teachers who appear to be spending excessive amounts of time in preparation for testing or engaging in item-teaching. Administrators must also be clear about the negative implications for students when discussing this issue during staff meetings. Ultimately, administrators should strive to promote assessment literacy in their schools. Assessment literate teachers understand how to use assessment to maximize student motivation and learning (Stiggins, 2000). These teachers are better situated to address the challenges of standardized testing and positively convey this challenge to their students.

Administrators and school district personnel also need to be skeptical of results that seem outside the norm for individual students and schools. Is it reasonable to assume a child scored at such a high level on a standardized test when they consistently get poor marks in class for this subject? Similarly, is it reasonable to assume a school's performance improved at such a high rate over such a short period of time? A school that consistently scores in the lower quartile (i.e., below 25<sup>th</sup> percentile) is unlikely to jump to the top quartile (i.e., above 75<sup>th</sup> percentile) in one or two years. District personnel should visit these schools to ascertain if inappropriate test preparation methods are being utilized. If this is the case, additional professional development assistance should be provided. In the fervor to bolster performance, administrators should not lose sight of the primary objective of these tests: providing an accurate measure of student performance. Those who allow teaching to the test techniques within their schools fail their students as well as their communities.

Depending on the jurisdiction, there may be a large degree of incongruity between the curriculum and the content represented by a standardized test. In such instances, teachers may be especially at-risk to abandon the mandated

curriculum in an effort to teach test items. Forcing teachers to choose between competing content leads to frustration and anxiety with testing procedures. Since its inception in 1997, Ontario's EQAO has ensured that teachers are principally responsible for the development, administration and scoring of the annual standardized assessments. Other jurisdictions would be wise to follow the lead set by this organization so that standardized tests are more closely aligned with prescribed provincial curriculum.

Teachers also need to have a clear description of the knowledge and skills represented by the test items in order to provide appropriate curriculum teaching (Popham, 2001). As a result, policy makers need to provide the resources for professional development training prior to the administration and reporting of results. Too often the cart comes before the horse, and teachers are forced to play catch up with respect to educational reform initiatives such as the implementation of high-stakes standardized testing. Behuniak (2002) suggests that administrators and teachers should be provided with support materials and workshops at least two years before they are held accountable – and farther in advance for truly high-stakes assessments.

Policy makers also need to consider that assessment of student progress should never be dependant on a single measure. Large-scale standardized tests are but one tool to assist in student development and learning. Such tests by themselves cannot produce the desired improvement to schools, because the tests do not deal with matters of teacher effectiveness or student motivation (Stiggins, 1999). Emphasizing standardized testing and formal evaluations has a negative impact on intrinsic motivation (Miller & Tovey, 1996). Research suggests that some students are likely to utilize inappropriate strategies for test preparation and writing as a result of their decreased motivation for test taking (Burger & Krueger, 2003). Examples include drawing pictures or writing editorial comments on the relevance of the test instead of completing the questions.

Emphasizing standardized testing also skews educators and the public's attention from what we want children to learn rather than what we can easily measure. For example, most experts agree that a balanced literacy curriculum should include components related to reading, writing, as well as speaking and listening (Sampson, Rasinski, & Sampson, 2003). Unfortunately, only reading and writing are typically measured by standardized school tests. Does this mean that speaking and listening are not valuable or worthy of attention within the curriculum? Obviously students need to utilize their speaking and listening skills to be successful in schools and higher education settings. Focusing heavily on standardized testing can serve as a distraction from more meaningful debates on the knowledge, skills, and dispositions we seek in school children.

Furthermore, policy makers should acknowledge that the practice of ranking schools is detrimental to schools and districts and take steps to eliminate this practice. Consider the advice provided by testing agencies such as EQAO, regarding the reporting of results.

The results of these assessments should not be use to rank schools or boards, for that would entail reducing the results to a single score or number. With the comprehensive nature of the data provided to schools and boards, ranking would be misleading. Ranking does not contribute to the well-being of Ontario students (EQAO, 1998, p.15).

Pitting schools and districts against one another only serves to facilitate the adoption of maladaptive test preparation which adversely affects student learning.

Authentic standards-based school improvement can be achieved without relying heavily on quantitative measures of student performance (Thompson, 2001). Multiple measures, including a variety of formats such as writing, open-ended response questions, and performance-based tasks provide a more balanced assessment approach that can positively impact student knowledge and skills (Jones, 2004). The increase in relevance of other indicators of student performance should also lead to a corresponding decrease in teaching to the test techniques. In many respects, the utilization of multiple measures of student performance is essential to preserving the validity of standardized tests that policy makers hold in such high regard.

Lastly, the public needs to become critical consumers of educational data. A child's education cannot be reduced to a single set of numbers reported in a newspaper or website. The simplistic idea that achievement scores in themselves are adequate and sufficient indicators of school performance is naive and should be dispelled (Earl, 1998). If parents want to infer how well their children will do in school next year, they need to make inferences about the broader domain and not about the specific objectives that are tested on a particular standardized test (Mehrens, 1989). Being knowledgeable concerning the types of conclusions that can and cannot be drawn from standardized test batteries seems essential. Overall, parents need to clearly understand the purpose of the tests that are being imposed on their children.

### **Conclusion**

Province-wide achievement testing of elementary and high school students is now the norm throughout Canada (Canadian Federation of Teachers, 1999). Unfortunately, the increased salience of these assessment measures has led some teachers to adopt maladaptive test preparation strategies such as teaching to the test. This practice of teaching to the test rarely affects learning and has a detrimental effect on the teaching profession as a whole. Training in appropriate

test preparation activities provides educators with the important skills they require to be effective teachers. Curriculum teaching provides students, educators, policy makers and the public with a sound basis for making future decisions. Most importantly, students will be able to engage in authentic learning that promotes future success.

### References

- Behuniak, P. (2002). Consumer-referenced testing. *Phi Delta Kappan*, 84(3), 199-207.
- Borich, G. D., & Tombari, M. L. (2004). *Educational assessment for the elementary and middle school classroom* (2<sup>nd</sup> ed.). New Jersey: Pearson Education Inc.
- Burger, J. M., & Krueger, M. (2003). A balanced approach to high-stakes achievement testing: An analysis of the literature with policy implications. *International Electronic Journal for Leadership in Learning*, 7(4). Online at <http://www.ucalgary.ca/~iejll>.
- Canadian Federation of Teachers. (1999). *Province-wide assessment programs*. Online at <http://www.ctf-fce.ca/e/what/other/assessment/testing-main.htm>.
- Chatterji, M. (2003). *Educational assessment*. Boston: Allyn and Bacon.
- Earl, L. (1998). Developing indicators: The call for accountability. *Policy Options*, 6, 20-25.
- Earl, L., Levin, B., Leithwood, K., Fullan, M., Watson, N., Torrance, N., Jantzi, D., Mascall, B., & Volante, L. (2003). *England's National Literacy and Numeracy Strategies: Final report of the external evaluation of the implementation of the strategies*. Department of Education and Employment, England.
- EQAO. (1998). *Educators handbook*. Toronto, Ontario: Queen's Printer of Ontario.
- Goodnough, A. (1999, December 9). New York City teachers nabbed in school-test cheating scandal. *National Post*, p. B1.
- Gronlund, N. E. (2003). *Assessment of student achievement* (7<sup>th</sup> ed.). Boston: Allyn and Bacon.
- Hart, P., & Teeter, R. (2001). *A measured response: Americans speak on education reform*. Princeton, NJ: Educational Testing Service.
- Herman, J. L. (1992). What research tells us about good assessment. *Educational Leadership*, 49(8), 74-78.
- Jones, K. (2004). A balanced school accountability model: An alternative to high-stakes testing. *Phi Delta Kappan*, 85(8), 584-590.
- Levinson, C. Y. (2000). Student assessment in eight counties. *Educational Leadership*, 57(5), 58-61.
- Linn, R. (2000). Assessment and accountability. *Educational Researcher*, 29(2), 4-17.
- MacDonald, T. (2001). To test or not to test: A question of accountability in Canadian Schools. Online at <http://policy.ca/archive/20010622.php3>.
- McMillan, J. H. (2000). Fundamental assessment principles for teachers and school administrators. *Practical Assessment, Research & Evaluation*, 7(8). Online at <http://pareonline.net/getvn.asp?v=7&n=8>.
- Mehrens, W. A. (1989). Preparing students to take standardized achievement tests. *Practical Assessment, Research & Evaluation*, 1(11). Online at <http://pareonline.net/getvn.asp?v=1&n=11>.
- Mehrens, W. A., & Kaminski, J. (1989). Methods for improving standardized test scores: Fruitful, fruitless, or fraudulent? *Educational Measurement: Issues and Practices*, 8(1), 14-22.
- Miller, E., & Tovey, R. (Eds.). (1996). *Motivation, achievement, and testing*. Boston: Harvard Education Press.
- Mitchell, K. (1997). What happens when school reform and accountability testing meet? *Theory into Practice*, 36(4), 262-265.
- Neil, M. (2003a). High stakes, high risk: The dangerous consequences of high-stakes testing. *American School Board*

*Journal*, 190(2), 18-21.

Neil, M. (2003b). The dangers of testing. *Educational Leadership*, 60(5), 43-46.

Payne, D. A. (2003). *Applied Educational Assessment* (2<sup>nd</sup> ed.). Toronto: Wadsworth Group.

Popham, W. J. (2001). Teaching to the test. *Educational Leadership*, 58(6), 16-20.

Popham, W. J. (2000). The mismeasurement of educational quality. *School Administrator*, 57(11), 12-15.

Sacks, P. (2000). Predictable losers in testing schemes. *School Administrator*, 57(11), 6-9.

Sampson, M. B., Rasinski, T. V., & Sampson, M. (2003). *Total Literacy* (3<sup>rd</sup> ed.). Canada: Wadsworth.

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.

Simner, M. L. (2000). *A joint position statement by the Canadian Psychological Association and the Canadian Association of School Psychologists on the Canadian press coverage of the province-wide achievement test results*. Online at [http://www.cpa.ca/documents/joint\\_position.html](http://www.cpa.ca/documents/joint_position.html).

Smith, M. L., & Fey, P. (2000). Validity and accountability of high-stakes testing. *Journal of Teacher Education*, 51(5), 334-344.

Stiggins, R. (2000). Learning teams for assessment literacy. *Orbit*, 30(4), 43-46.

Stiggins, R. (1999). Assessment, student confidence, and school success. *Phi Delta Kappan*, 81(3), 191-198.

Taylor, A. R., & Tubianosa, T. (2001). *Student assessment in Canada: Improving the learning environment through effective evaluation*. Kelowna, BC: Society for the Advancement of Excellence in Education.

Thompson, S. (2001). The authentic standards movement and its evil twin. *Phi Delta Kappan*, 82(5), 358-362.

Wright, W. E. (2002). The effects of high stakes testing in an inner city elementary school: The curriculum, the teachers, and the English language learners. *Current Issues in Education*, 5(5). Online at <http://cie.ed.asu.edu/volume5/number5>.