



Standardized Tests: Bellwether of Achievement?

Matt Deerman

Jacksonville State University

Cherith Fluker

Jacksonville State University

Elizabeth Panik

Jacksonville State University

Jimmy Powell

Jacksonville State University

Kathy Shelton

Jacksonville State University

Carol S. Uline

Jacksonville State University

Charles E. Notar

Jacksonville State University

700 Pelham Road N

Ramona Wood A3

Jacksonville, Alabama 36265-1602

Tel: 1-256-782-5959 E-mail: cnotar@jsu.edu

Abstract

This article constitutes the culmination in a review of 53 sources by a group of first year master's degree students in a problems and issues in secondary education course to determine "Standardized Tests: Bellwether of Achievement?" Standardized tests are used for deciding whether or not students are achieving academically in American schools. Is this the right mechanism? Are they fair for all students? Is there a better way? There are pros and cons to standardized testing, and these will be discussed in the paper.

Keywords: Academic achievement, Accountability Assessment, High stakes testing, Standardized tests, Standards, Test and measurement

Standardized testing is a popular method of measurement in modern American schools. High expectations raise the bar for student achievement. Tests can provide critical benchmarks of students' knowledge and they can provide validation for the quality of the nation's public schools.

Emphasis on standardized testing helps to answer the call for school reform. Standardized tests are among the most objective ways to measure student performance. Electronic scanning makes them easy to score. Standardized tests contain validity and reliability standards.

Strict standards for administration of standardized tests authenticate objectivity and reliability. They are administered with a standard set of instructions which are read aloud, word for word, to all test takers. Standardized tests have set time limits and are intended to be administered in similar testing conditions. Test results are compared to a sample group or norm. Statistical methods determine placement of the test scores on a normal curve, which when graphed, appears as a bell-shaped curve. Scores plotted on the bell-curve may be used to compare students with local, state, and

national levels of test takers (Moriarty, 2002). Favorable standardized test results from schools provide confidence to a nation in competition with other countries engaged in global economics.

1. Historical Background of Standardized Tests in American Schools

Standardized testing began in the United States as early as the 1900s to determine a student's individual intelligence quotient. Prior to 1965, standardized tests were not used in the early grades. These years were considered a time in a child's life for natural growth and development. However, with the Soviet Union's launching of Sputnik in 1957, American schools pushed for higher achievement in science and mathematics. With the federal government providing resources to help fund schools in the 1960s, evaluation demands grew. Standardized tests met the accountability requirements for federal funding. They were relatively inexpensive and easy to administer. In the 1970s, accountability for federal funding became more stringent. Eventually, standardized tests became the definers of standards in all subject areas. By the 1980s, testing of young children became commonplace. Today, standardized tests are used educationally from kindergarten through college.

2. Uses of Standardized Tests

When people think about the uses of standardized tests, they generally consider college admissions. Defenders of the use of standardized tests for college admissions assert that the tests "lend predictive power in the admissions process" (Calvin, 2000).

Of course, there are other reliable predictors of success in college, such as grades earned during the junior and senior years of high school. Notwithstanding, standardized tests continue to be used as an objective measure for college admittance as well as for many other purposes.

An entire battery of standardized testing is used to distinguish special needs students from regular students. As a diagnostic tool, educators, counselors, and other professionals plan Individualized Education Programs (IEPs) for struggling students based on the results of the norm-referenced tests. Such a battery of tests can "point out specific misconceptions or problem areas that are hindering progress" for specific students who obviously learn differently from those students who succeed with the traditional methods of listening, reading, writing, and testing (Fremer & Wall, 2003).

Another use of standardized tests is for counseling services. Students exhibiting unusually high levels of stress or who engage in substance abuse may be recommended for testing. Test results may suggest counseling aimed at preventing harmful or dangerous behaviors (Fremer & Wall, 2003). Counselors may also use standardized tests for purposes of student self discovery. Results of attitude surveys and interest inventories help students choose careers.

For accountability purposes, standardized achievement tests are used to determine academic knowledge and skills. Such tests may be used to place students in remedial, regular, or advanced classes. High schools began using exit examinations in the late 1970s as a means of insuring that students meet minimum state requirements for graduation. This more exclusive use of standardized tests led educational institutions to the current practice of high-stakes testing.

Today's emphasis on high-stakes testing has been described as "the learning through standards and accountability era of American education" (Sloane & Kelly, 2003).

Testing definitely plays an important role in maintaining accountability in the school systems. Objective measures are necessary to determine if students are mastering the necessary skills and critical concepts needed to ensure future success. But the teacher and school accountability factor raises issues.

With the recent legislation of No Child Left Behind (NCLB), more and more emphasis is placed on accountability and teacher evaluation. This legislation requires regular accountability testing as a prerequisite for receiving and continuing to receive federal funds (Fremer & Wall, 2003). As a result, teachers whose students fail to show adequate progress, must engage in professional development workshops to improve their teaching techniques. Administrators and subject matter experts work with teachers to help them become more successful in the classroom.

3. Pros and Cons of High-Stakes Testing

Questions continue to surface about the benefits of over reliance on standardized test results. Is standardized testing the most beneficial strategy for measuring success of both students and teachers? Should a single standardized instrument be used to determine promotion and/or graduation? Are standardized tests fair to all students?

In recent years, both teachers and students have been held accountable for results of standardized testing. Oftentimes, the quality of a school is based on how well students do on standardized tests. Both teachers and students feel the pressures of administrators, peers, and parents to perform well on standardized tests (Abrams, Pedulla, & Madaus, 2003). The pressure on teachers can lead to irregularities in the administering of tests which, of course, can skew the results (Gay, 1990). Many teachers have admitted to teaching to the test. They adjust their objectives to teach

specific content that the test covers. If test scores go up, can these improvements be trusted when teachers and administrators are held accountable for the test scores?

When teachers are evaluated solely on how well their students perform on a standardized test, they abandon creative, well-rounded, interdisciplinary teaching with accommodation for diverse learning styles. Teachers give up using strategies they consider most productive to academic achievement, such as hands-on projects, inquiry problem solving activities, and scenarios which engage critical thinking skills. Instead, they focus on rote memorization and rehearse students in the art of bubbling in answers on practice tests.

Not only does high-stakes standardized testing negatively affect the art of teaching, it negatively affects student learning and self esteem. The pressure on students to perform well on a quantitative measurement of learning at one specific point in time negates the qualitative progress a student makes over a period of time. All students learn differently and perform differently on different types of assessments. Writing samples, oral presentations, and research projects are examples of assessment not measured by standardized tests.

Standardized tests are unfair not only in terms of non-accommodation for diverse learning styles, but they do not take into account subgroups, such as racial and ethnic minorities, students with disabilities, students with limited language proficiency, and students from low socioeconomic groups (Bracey, 2001). Test content and testing conditions tend to ignore the issue of diversity. Some schools teach "testwiseness" to help bridge the diversity disparity. Students need to know how to take a standardized test. Awareness of certain cognitive, secondary cues, present in multiple-choice test items can help elevate test results. One strategy is to try to answer the question in one's head before reading the possible answers. This practice lessens confusion by the answer choices. Other strategies include "organization, planning, and time management" as vital skills in test taking (Loulou, 1995).

One of the goals of standardized testing is to compare schools on a national level. Every state has the same content covered on each of the tests, but the level of difficulty is different from state to state (Linn & Kiplinger, 1995). Since all state tests are not the same, and there is no national test in place, it is difficult to compare achievement among schools throughout the United States.

4. Alternatives to Standardized Testing

Alternative methods of assessment have shown to have some positive effects for students that may outweigh the advantages of standardized testing (Seeley, 1994). One alternative method is performance assessment where students demonstrate a skill, based on a specific behavioral objective. Another assessment alternative is maintaining portfolios where students select their own work samples. These student products plainly reveal learning progress to the student, the teacher, and the parents, who traditionally ask to see their children's work. Journals and interviews are helpful assessment tools. Attitude inventories and opinion surveys are other positive methods for alternative assessment (Travis, 1996).

Teachers like the versatility and creativity alternative methods allow. Teachers may use these assessment activities often, even daily if desired (Bol, Ross, Nunnery, & Alberg, 2002). Teachers may use these measures to grade all students, even those with exceptionalities. Creative teachers may customize alternative methods to adapt assessment activities to the specific needs of diverse learners.

Students tend to prefer alternative assessment. They feel more in control of their learning environment. Students crave change and flexibility in learning and assessment activities. Alternative assessment activities motivate students to rise to the challenge and try harder to excel (Allen & Flippo, 2002).

However, there are negative aspects of alternative assessment. The greatest drawback teachers find with alternatives is converting the assessment data into letter grades and percentages. Teachers also worry that alternative grading procedures may not accurately cover material in the mandated standardized tests. Certainly, alternative methods require more time and effort than standardized testing. For these reasons, teachers, especially less experienced teachers, may be reluctant to try alternatives for assessing student achievement.

Education cannot be a "laissez-faire" practice; accountability is necessary. Common sense suggests a combination of standardized testing and alternative assessment methods. Sole reliance on high-stakes testing tends to defeat students, demean teachers, and discredit schools. Testing is an extremely valuable part of educational assessment, but it is only part of the formula for quality learning. Additional research may help to delve more deeply into the intended and unintended consequences of standardized testing in educational decision making.

References

- Abrams, L., Pedulla, J., & Madaus, G. (2003). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory Into Practice, 42*(1), 18-29
- Allen, D., & Flippo, R. (2002) Alternative assessment in the preparation of literacy educators: Responses from students. *Reading Psychology, 23*(2), 15-26.

- Baker, E., O'Neil, H., & Linn, R. (1993). Policy and validity prospects for performance-based assessment. *The American Psychologist*, 48(12), 1210-1218.
- Baker, E. (2001). Testing and assessment: A progress report. *Educational Assessment*, 7(1), 1-12.
- Boaler, J. (2003). When learning no longer matters: Standardized testing and the creation of inequality. *Phi Delta Kappan*, 84(7), 502-506.
- Bol, L., Ross, S., Nunnery, J., & Alberg, M. (2002). A comparison of teacher's assessment practices in school restructuring models by year of implementation. *Journal of Education for Students Placed at Risk*, 7(4), 407-423.
- Borja, R. R. (2007). Nebraska swims hard against testing's tides *Education Week*, 26(24), 32-34.
- Bracey, G. (1996). Money improves test scores—even state-level SATs. *Phi Delta Kappan*, 78(1), 91-93.
- Bracey, G. (2001). Test scores in the long run. *Phi Delta Kappan*, 82(8), 637-638.
- Calvin, A. (2000). Use of standardized tests in admissions in postsecondary institutions of higher education. *Psychology, Public Policy, and Law*, 6(1), 20-32.
- Causey-Bush, T. (2005). Keep your eye on Texas and California: A look at testing, school reform, No Child Left Behind, and implications for students of Color. *Journal of Negro Education*, 74(4), 332-343.
- Cheek, J., Bradley, L., Reynolds, J., & Coy, D. (2002). An intervention for helping elementary students reduce test anxiety. *Professional School Counseling*, 6(2), 162-164.
- Chudowsky, N., & Pellegrino, J. (2003). Large-scale assessments that support learning: What will it take? *Theory Into Practice*, 42(1), 75-83.
- Educational Testing Service. (2001). A measured response: Americans speak on education reform. Retrieved May 14, 2004, from the Educational Testing Service Web site: <http://www.ets.org/aboutets/americaspeaks/survey2001.html>
- Fremer, J., & Wall, J. (2003). Why use tests and assessment? North Carolina: Counseling and Student Services. (ERIC Document Reproduction Service No. ED480036)
- Gay, G. (1990). Standardized tests: Irregularities in administering of tests which affect test results. *Educational Tests & Measurements*, 17(2), 93-103.
- Geiger, M. (1997) An examination of the relationship between answer changing, testwiseness, and examination performance. *The Journal of Experimental Education*, 66(1), 49-60.
- Gloe, D. (1999). Study habits and test-taking tips. *Dermatology Nursing*, 11(6), 439-447.
- Goldberg, M. (2005). Losing students to high-stakes testing. *Education Digest*, 70(7), 10-19.
- Gronna, S., Jenkins, A., & Chin-Chance, S. (1998). The performance of students with disabilities in a norm-referenced, statewide standardized testing program. *Journal of Learning Disabilities*, 31, 482-493.
- Harmon, M., Morse, D., & Morse, L. (1996). Confirmatory factor analysis of the Gibb experimental test of testwiseness. *Educational and Psychological Measurement*, 56, 276-286.
- Herman, J., Abedi, J., & Golan, S. (1994). Assessing the effects of standardized testing on schools. *Educational and Psychological Measurement*, 54, 471-482.
- Herman, J., Klein, D., & Wakai, S. (1997). American student's performance on alternative assessment: Do they know it's different? *Assessment in Education: Principles, Policy, & Practice*, 4(3), 339-349.
- Holloway, J. H. (2001). The use and misuse of standardized tests. *Educational Leadership*, 59(1), 77-78.
- Karantonis, A., & Sireci, S. G. (2006). The Bookmark Standard-Setting Method: A literature review. *Educational Measurement: Issues & Practice*, 25(1), 4-12.
- Linn, R. (2001). A century of standardized testing: Controversies and pendulum swings. *Educational Assessment*, 7(1), 29.
- Linn, R., & Kiplinger, V. (1995). Linking statewide tests to the national assessment of educational progress: Stability of results. *Applied Measurement in Education*, 8(2), 135-155.
- Loulou, D. (1995). *Making the A: How to study for tests*. Washington, D.C.: Office of Educational Research and Evaluation. (ERIC Document Reproduction Services No. ED385613)
- Mabry, L. (1999). Writing to the rubric. *Phi Delta Kappan*, 80(9), 673-679.
- McCabe, P. (2003). Enhancing self-efficacy for high-stakes reading tests. *The Reading Teacher*, 57(1), 12-20.

- McCaffrey, D., Hamilton, L., & Stecher, B. (2001). Interactions among instructional practices, curriculum, and student achievement: the case of standards-based high school mathematics. *Journal for Research in Mathematics Education*, 32(5), 493-517.
- Moriarty, F. (2002). *History of standardized testing*. Oct. 4, 2007 from www.or.assortment.com.
- Norwood, K. (1994). The effect of instructional approach on mathematics anxiety and achievement. *School Science and Mathematics*, 94(5), 248-254.
- Office of Educational Research and Improvement. (1995). *Cost of a National Examination*. ERIC Clearinghouse on Assessment and Evaluation, Washington, DC. (ED385611)
- Popham, W. (2003). The seductive allure of data. *Educational Leadership*, 60(5), 48-51.
- Popham, W. J. (2006). Assessment for learning: An endangered species? *Educational Leadership*, 63(5), 82-83.
- Priestley, M. (2000). 10 tips for higher test scores. *Instructor*, 110 (3), 30-32.
- Rescorla, L., & Rosenthal, A. (2004). Growth in standardized ability and achievement test scores from third to tenth grade. *The Journal of Educational Psychology*, 96(1), 85-96.
- Ritter, S., & Idol-Maestas, L. (1986). Teaching middle school students to use a test-taking strategy. *Journal of Educational Research*, 79(6), 350-357.
- Salpeter, J. (2007). Is more testing the answer? *Technology & Learning*, 27(10), 18-23.
- Samuelson, S. (2001). Student testing: The stakes are rising. *State Legislatures*, 27(8), 30-34.
- Scherer, M. (2005). Reclaiming testing. *Educational Leadership*, 63(3), 9.
- Seeley, M. (1994). The mismatch between assessment and grading. *Educational Leadership*, 52(2), 4-6.
- Sloan, F., & Kelly, A. (2003). Issues in high-stakes testing programs. *Theory Into Practice*, 42(1), 12-17.
- Smith, M. S. (2005). Improving student achievement by measuring ability, not content. *T.H.E. Journal*, 32(9), 56.
- Sgoutas-Emch, S., & Johnson, C. (1998). Is journal writing an effective method of reducing anxiety towards statistics? *Journal of Instructional Psychology*, 25(1), 49-58.
- Travis, J. (1996). Meaningful assessment. *Clearing House*, 69(5), 308-313.
- Vito, P. (1991). *On standardized testing*. Urbana, IL: Clearinghouse on Elementary and Early Childhood Education (ERIC Document Reproduction Service No. ED338445)
- Walen, S., & Hirstein, J. (1995). Classroom vignette: An alternative-assessment tool. *Teaching Children Mathematics*, 1(6), 362-366.
- Wesson, K. (2001). The "volvo effect"- questioning standardized tests. *Young Children*, 56(2), 16-18.
- Williams, N. (2003). Thinking outside the bubble. *Educational Leadership*, 61(3), 82-83. Wightman, L. (2000). The role of standardized admission tests in the debate about merit, academic standards, and affirmative action. *Psychology, Public Policy, & Law*, 6(1), 90-100.
- Wildemuth, B. (1984). *Alternatives to standardized tests*. Washington, D.C.: National Institute of Education (ERIC Document Reproduction Service No. ED286938)