

THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF CURRICULUM & INSTRUCTION

STANDARDIZED TESTING IN THE 1960S AND 1970S: EXPLORING THE IMPACT ON
TODAY'S EDUCATIONAL CLIMATE

EMMA E. PEBLEY

Fall 2010

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Secondary Education
with honors in Curriculum & Instruction

Reviewed and approved* by the following:

Dr. David Gamson
Associate Professor of Education
Thesis Supervisor

Dr. Anne Whitney
Assistant Professor of Education
Honors Adviser

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

This paper serves as an exploration of the application of standardized testing in the 1960s and 1970s. In the current fervor for standardized tests, it is important to examine the history of these tests. By reviewing what types of tests existed in the past, along with the contemporary literature that helps to explain the feeling about these tests, I hope to present a picture of this period in standardized testing history. It is important to consider the tests themselves, the literature about the tests, and also the social and political contexts surrounding them. This discussion of the literature and testing of the past provides a basis from which to examine the testing situation of today's education and warrants an examination of what, if, and how much our ideas and applications of testing have changed.

TABLE OF CONTENTS

Abstract.....	p. i
Standardized testing in the 1960s and 1970s: Exploring the impact on today's educational climate.....	p. 1
The 1960s.....	p. 5
The 1970s.....	p. 12
The Impact on Our Present Situation: From the Perspective of a Student and Teacher.....	p. 16
References.....	p. 21
Figures.....	p. 24
Academic Vita	p. 26

Standards and testing have been a part of the American educational mind since the inception of public schools. It has always been important to be able to determine whether students have learned what teachers have taught. It is important to determine whether what teachers are teaching is, in fact, what students need to know. It used to be that the purpose of schooling was to master the “three R’s”: Reading, Writing, and ‘Rithmetic. If, at the end of his or her schooling, a student could successfully do those three things, his/her education had met the standards of the time.

But as education has evolved, so have the standards and tests that are applied in the public school system. Standardization is a term first coined in 1873, and it is still used to mean to compare with or test by a standard, with a standard being “something set up by authority as a rule for the measure of ...value or quality” (“Webster’s,” 2001). The value or quality of interest in education has always been linked to student learning, measuring each student’s results to a determined level of proficiency. It may not always have been a specific test, but even rote memorization served as a standard in its time. Students were expected to complete their recitations to a certain standard, with that standard of completion and correctness indicating success or failure. While standards and standardization have not always been buzzwords, they have always been present in our system of education.

Testing has also been a part of the educational consciousness in some form or another. Teachers applied tests to determine the degree to which a student had learned a given set of information, applicable only to the information taught in that single classroom. These tests are still in use today, but as teacher and curriculum variation proliferated, it became difficult to determine whether two students from different contexts: different schools, different states, different home environments, etc. were prepared by their education to the same level of

proficiency. Testing that went beyond a single classroom became particularly necessary as colleges and universities began to proliferate after World War II (“Secrets of the SAT,” n.d). Some combination of testing and standards that was easily transportable to different schools throughout the country and represented the necessary skills for entrance into college was necessary, and development of these tests began.

Intelligence testing was already breaking down these barriers, with the SAT and the Army-Navy College Qualification Test being used to determine intelligence and scholarship eligibility in large numbers of individuals across the country at one time. In 1943, the Army-Navy College Qualification Test was administered to 316,000 high school seniors. This proved that the same multiple-choice test could be administered to an enormous group (“Secrets of the SAT,” n.d.). The precedent was set.

While not linked directly to a single curriculum, these tests had a purpose: to compare the intelligence of vast numbers of people in hopes of categorizing individuals so as to better determine their merit. The tests were helping to simplify the process of determining intelligence; not only could an individual’s intelligence be immediately scored, but thousands of individual’s scores could be tabulated and categorized at the same time (Lehmann, 1999).

The terms “standardized” and “testing” were first combined in the 1950s, following the passage of the National Defense Education Act (NDEA). Standardized tests became important in determining whether American students were falling farther behind their international counterparts in school subjects, particularly math and science. These tests were the first to be made standardized, that is, the exact same test could be administered to students all over the country and even all over the world.

Research on the specific testing effects of NDEA is not extensive, but many researchers and writers have focused on the beginnings of these private testing companies or on the funding, implementation, and impact of NDEA on the education community as a whole. Most recently, Wayne Urban explains in some detail in his speech to the Midwest History of Education Society the implications of Title V of the National Defense Education Act. He states simply, “Title V was for testing...for the purpose of identifying and encouraging able students” (Urban, 2008). The title authorized \$15,000,000 per year for states to create programs of testing and counseling for students. He also addresses another side of the education industry that was boosted by the passage and subsequent funding of NDEA, the United States Office of Education (OE). NDEA granted OE the power to become an influential federal agency, and, despite its name change to the Department of Education, the OE is still a powerful voice today. It was NDEA that gave the OE footing, granting it funding for more jobs and projects, many of which involved overseeing NDEA Titles.

Gail Kelly and Maxine Sellers wrote a history of educational reform in New York State in 1985. They observed that NDEA legislation did fund the educational aims of increasing the amount of science instruction equipment in schools, but the testing that was used as its measure of success effectively “undermined academic standards and promoted solely test-taking skills” (Kelly & Seller, 1985). Their study spans the history of New York State, and notes each time federal or state legislation was introduced to the New York education system, testing resulted.

Robert Linn focused more generally on the initiation of standardized testing in schools and the reason for its persistence. In “Educational Testing and Assessment: Research Needs and Policy Issues,” published in *American Psychologist* in 1986, he maintains that standardized testing grabbed hold of America’s public education system after its mandated beginnings

because of its promise of efficiency. A good test allowed administrators to track students and predict outcomes. The other major reason for their success, he argues, is that they are coupled, by the publisher, with a curriculum designed to help improve student scores on those tests (Linn, 1986).

Peter Airasian discusses the effects of government intervention on public education in his 1987 article “State Mandated Testing and Educational Reform: Context and Consequences.” He notes that the politicalization of education that resulted from NDEA and the success of Sputnik necessitated testing. Government, he concludes, needed ways to ensure that their money in schools was being used well, and that their goals were being reached. Testing becomes the ideal vehicle for both aims because, as Airasian discusses, adults trust their objectivity and ability to reveal the truth. “Many different types of tests were utilized to carry out the desired monitoring, including traditional school-based standardized achievement and ability tests, the Scholastic Aptitude Tests, and new state assessment tests that proliferated at this time,” Airasian writes, plainly drawing the relationship between federal regulation and the testing industry (Airasian, 1987).

In 2006, the Institute for Defense Analyses’ Science and Technology Policy Institute published a paper that detailed the funding of NDEA and the effects of that funding. Of particular interest is their assertion that the number of standardized tests administered during the tenure of NDEA and the number of students taking the tests ballooned.: “The number of standardized aptitude tests given ...dramatically increased—from approximately one-third of K–12 students in 1958 to nearly all students by 1966. During those years, the number of standardized aptitude tests administered ... increased from 10 million (2 million of which were paid for under the NDEA) to 45 million (9 million of which were paid for under the NDEA)”

(Flatteau, Bracken, Van Atta, Bandeh-Ahmadi, de la Cruz, & Sullivan, 2006). In eight years then, the number of students taking these standard examinations tripled, and the number of tests administered doubled *twice* and kept increasing. These tests were funded through the act because of their importance in gauging the progress of America's progress in science education, and as testing kept giving results, schools kept giving tests.

Each of these publications serves to illustrate that the effects of NDEA and the federal funding of education have been of interest to researchers for thirty years. Each publication agrees that NDEA was a new kind of policy, a new kind of oversight for education, and a new kind of testing for America. They also have the common aim of understanding NDEA as it fits into the contemporary influence, of each piece, of education in the United States.

This incredible proliferation of testing and its funding brought the United States into the 1960s. Aptitude tests, as well as curriculum-based tests, were on the rise, with schools and administrators eager to prove progress just as testing companies were eager to prove that their combined tests and curricula could improve the education of all students.

Testing in the 1960s

As education barreled headfirst into the 1960s, the prevailing feelings in education were of fear. Were we falling behind in education? Could the nation that saved the world from certain destruction in WWII truly be losing its foothold as a superpower already? Money was pouring into curricula and test creation companies at an astronomical rate, all with the promise of more efficient and effective methods of teaching students and testing their recollection of what was taught. These feelings and questions took education into an "era" commonly referred to as "Scholarly Structuralism."

The period of Scholarly Structuralism, beginning in 1957 with the launch of *Sputnik*, was a period when the unrest and wars of the previous era continued gathering strength. Massive allocations of federal aid into education, a result of urging by Congress and President Johnson, resulted in massive changes in testing and curriculum theory. The tests facilitated by Title V of NDEA created an importance in curriculum of the *structure* of disciplines, focusing on using school time efficiently to ensure that students learned the processes and skills that could be used efficiently across disciplines (Glatthorn, Boschee, & Whitehead, 1999). This type of education lent itself to widespread standardized testing because processes are uniform across classroom, district, and state lines. If the way of teaching something was the same everywhere, the results should differ only in the amount that the students understood about the suggested materials.

In his article “The Achievement Decline of the 1970s,” Paul Copperman notes that from 1958 to 1964, the Stanford Achievement Tests (SAT) showed a steady increase in reading and math skills in the grades tested (Copperman, 1979). He also notes the names and publishers of several tests that grew to prominence during the NDEA period, noting that re-norming, or establishing a new mean score for the exams, was required for most of them as the number of students taking the tests skyrocketed. The tests he discusses include the Metropolitan Achievement Tests, the Stanford Achievement Tests published by Harcourt, the California Achievement Tests published by McGraw-Hill, the Iowa Tests of Basic Skills by Houghton Mifflin, and the Gates-McGinitie Reading Test, also a Houghton Mifflin publication (Copperman, 1979). Many of these tests still exist in some form or another, continuing to change as necessary in order to best assess the learning of students in public schools.

The Iowa Tests of Basic Skills (ITBS) is one test that greatly proliferated during this time period. This test came into being in 1935 as the Iowa Every Pupil Test, and has continued to

gain momentum in the state throughout the twentieth century (“ITBS,” n.d.). This test series provides students with results in different content areas, basing their results on a “monthly progress” meant to indicate where on the spectrum of the school year the individual’s skills fell. This was an interesting perspective in the 1960s. Students could be measured on the continuum of their abilities—smarter, more advanced students would have results that demonstrated a higher “month” of mastery—but they could also be measured in comparison with the curriculum—where in the year’s proposed instruction did the individual student’s skills fall. It also managed to adjust itself to the increasing number of students taking the test; it is still in use today in a similar form with almost every public high school in Iowa administering it at some point during a school year (Riverside, n.d.).

These tests are administered from kindergarten (level 5) to eighth grade (level 14) (Riverside, n.d.). Levels 5 and 6, made for kindergarten and first grade students, are untimed tests meant to engage students’ interest in the test and its questions, are read to the students by the teacher. The skill groups for these tests fall into the realms of math and reading, a trend that continues into the upper levels. Levels 7 and 8 of the test, for first and second grade students, are also orally administered to students, but more comprehensive subject testing is present. Social Studies and Science areas of testing, as well as a “sources of information” subtest, enter the picture (“ITBS,” n.d.). Levels 9-14 are timed tests, comprising 13 different tests, four in the Language category, and 3 in Mathematics. These tests were meant to assess student learning throughout the year, giving teachers an opportunity to assess students’ strengths and weaknesses and to adjust instruction accordingly. Directions for these tests were still given orally, and this pattern continues to this day.

Whether this ability to adjust instruction was being effectively used, however, was a question that many were unable to answer during the time period. The ITBS is a test that is still recommended to be given in the fall, so that these instructional adjustments can actually be made within a given school year, which separates it from most standardized tests of the time and even from standardized tests of today. Tests are typically given during the spring, when making instructional adjustments or individual plans is almost impossible due to the time left in the school year after the results are back (Winchester, 2006).

ITBS, because of its consistent use over the last eighty years, provides a very detailed picture about the use and effectiveness of this type of standardized testing over the decades of consideration. Achievement on this test increased steadily through the early and mid 1960s. This would indicate that the methods of instruction and of altering instruction to meet students' needs were effective in terms of the testing scores during this time period (Copperman, 1979). In the 1970s, this test will be revisited as evidence of the change in the climate of testing.

The biggest standardized test of the time was the Stanford Achievement Test. This test was created in 1926 to assess students at all levels of public schooling. This test was widely used during the 1960s and 1970s, and is still popular today. The test was a popular measure, in part, because of the increasing achievement of students on the test, even as renorming occurred. Between 1958 and 1964, eighth-grade test scores increased between 3.5 and 5% each year (Copperman, 1979). Achievement on this test began to decline in 1964, much to the confusion and chagrin of educators and test writers (Copperman, 1979). The rate of decline on the reading subtest was 3% per year, while in math the decline was approximately 7%, far outstripping the gains of the years previous.

An important publication during this period was the Conant Report. This document, written by James Bryant Conant, detailed what he felt would be the best curriculum for the American high school, making specific recommendations even for academically talented students. These recommendations, including 4 years of English, 1 year of science, and a course in “American problems,” were widely endorsed during the period because of the need for a new direction in education. The ease to which standardized testing could be applied to these recommendations even made them palatable to testing companies (Glatthorn, Boschee, & Whitehead, 1999). This period of time was marked by many examples of these “ideal” curricula, designed to simplify the process of schooling and of teaching – some were even referred to as “teacher-proof curricula” (Glatthorn, Boschee, & Whitehead, 1999).

Toward the late 1960s, these types of curricula, and their accompanying tests, began to fall out of favor. America was entering a violent time in its evolution, and as achievement declined on standardized tests designed to work with these specific plans, the plans began to change. The same kinds of changes were not necessarily witnessed in the tests themselves. Neither the Stanford Achievement Test nor the Iowa Test of Basic Skills, for example, underwent any major renovations, but the students to whom these tests were administered were pushing through a difficult age.

This era in education, spanning approximately from 1968 to 1974, can be classified as “Romantic Radicalism” (Glatthorn, Boschee, & Whitehead, 1999). Vietnam was in the headlines and the hearts of most Americans, and individual liberties and rights were making a public surge. Youth and volume were powerful players during this time period. Minority groups of all kinds found the strength and voice to assert their rights during this time period. The assassination of Dr. Martin Luther King, Jr., race riots, and desegregation efforts made for a

difficult educational climate, even with individuals like John W. Gardner pushing for legislation dealing with appropriate funding and resource allocation for all schools in the nation (“John Gardner,” n.d.) Education became a political machine, used as fodder for and against all of the social issues of the time. This period gave rise to a different kind of education, one less concerned with testing and quantifiable results and more concerned with rectifying the overt social and economic inequities in schools (Glatthorn, Boschee, & Whitehead, 1999).

The 1960s also witnessed another change in education legislation that had no small impact on this burgeoning age of testing. The Elementary and Secondary Education Act (ESEA) of 1965 represented the next step in federal influence over public education and greatly expanded its interests in the allocation of public funding for education. In 1965 alone, over \$1 billion were distributed to school districts, primarily under Title I funding (Thomas & Brady, 2005). While some of this money, as much as 15% in some studies, was misappropriated, the message was clear: the federal government needed to fund education where the resources were not up to par. Students needed to have access to the resources to be able to succeed.

Testing continued to be important in this era because of several factors. First, while Title I students could be identified by demographic information, including income, racial and ethnic background, and other relevant demographic and educational information, it was the achievement of Title I students that determined what part of their federal money allocation went where (Thomas & Brady, 2005). Title I became the new buzzword for education at this point, and the focus of testing began its shift from the achievement of all to the comparative achievement of the disadvantaged. These disadvantaged students included handicapped individuals, who were openly addressed in education for the first time during this period, as were students with limited English-language proficiency. ESEA forced schools to recognize and

assist these groups of students who had previously been marginalized by a lack of understanding, funding, and other teaching resources (Glatthorn, Boschee, & Whitehead, 1999). These pieces of the educational puzzle moved to the forefront, and the fervor over the positive impact of standardized testing began to patter out.

It is in this period as well that criticism of standardized testing as a primary means of assessment began. In 1965, for example, the interactional role of standardized testing began coming under fire. Why were educational papers about standardized testing and its application in classrooms gaining no notice? Why were standardized tests making no adjustments in terms of classroom application? The paper argued that it was the separation between the researchers and the writers of these tests and curricula that were resulting in the problems that were beginning to develop (Clark & Guba, 1965). These individuals argued for the integration of the research, development, and dissemination of these different testing practices, which they hoped would result in a more successful adoption phase, meaning simply that these tests would be relevant enough that districts and schools would want to use them as measures of standard achievement, rather than using them because of the ease of administration and the quantitative value of the results (Clark & Guba, 1965).

The 1960s ended in a period of unique turmoil. An American man was on the moon, American troops were entrenched in a foreign war, and education was facing racial, socio-economic, funding, and identity issues that would carry on into the next decade, setting the stage for a change of pace; testing would fade from its seat of primary importance.

Testing in the 1970s

While it is clear that testing began to lose its foothold as the primary topic for educational discussion as the nation headed into the seventies, it cannot be said that standardized testing stopped being used as a method of assessment in schools. Since its introduction, standardized testing has remained a popular form of assessment. The concern with standardized testing in the 1970s centered around an apparent bias against minority groups and on criticism of the institution of standardized testing in its entirety.

During this time period, the assertion that there was something inherently wrong with standardized tests and their usage was a fairly common one. Many groups condemned the use of these tests, while others insisted that these tests were valid measures of student achievement and asserted that those who took issue with the tests themselves instead had an issue with the application of the results. In this decade, the National Education Association (NEA) called for a “moratorium on all forms of standardized testing,” (Rudman, 1977). The American Federation of Teachers (AFT) took a different stance, arguing for more education on the strengths and weaknesses of tests and insisting that improvements to the tests that existed be made (Rudman, 1977).

Some teachers even expressed their frustration with standardized testing and its effect on the students. In *The English Journal*, Ernie Karsten published the following poem, titled “Standardized Testing—Reality Therapy?” to illustrate his impression of standardized testing’s impact:

Fourth period testing
wrestling/wresting
unknowns from knowns:
cubes/commas/cones
nicely neutral nouns
“sit/sat” frowns

opinion versus fact
multiply/divide/subtract.

Interest immediate
results remote
validity approximate
--no antidote? (1971)

The second stanza of the poem clearly indicates a frustration that was felt extensively during this time period. Districts wanted students taking these tests, creating the interest in them, and then the results from the tests were slow in returning, and especially in the case of spring testing, left no room to adjust instruction in a way that would be meaningful to students. Finally, depending on the test, whether it had been renormed, and what type of questions it was asking, the same student could earn a completely different result. Some educators felt that the entire process was flawed, and they wondered whether the tests themselves were necessary in any form.

Proponents of standardized testing did exist, however. Herbert Rudman, a professor of educational measurement whose research interests focused largely on standardized testing, actually took it upon himself to answer the eight most common criticisms about standardized testing. In an answer to the criticism that standardized testing determines what is taught in a classroom curriculum, he provided a graphic (included as Figure B on page #) that showed the number of steps that a standardized test goes through before it is published and used. The process, he stated, takes approximately eight to ten years, making it almost impossible for the test itself to determine what is being directly instructed in the classroom (Rudman, 1977). Additionally, Rudman answers a criticism about nationally normed test, claiming that these tests do not, in fact, lead to national comparisons that would change the way teachers in each state educate their students. He answers that, once again, a standardized test is not the determiner of

curriculum in any classroom, and as such, a nationally normed test could not lead to a set of national standards and curriculum (Rudman, 1977).

Others advocated neither for the unabated continuation of standardized testing nor for the immediate dismissal of standardized testing from schools. They advocated instead for a more responsive system of standardized testing, one that focused more directly on curriculum as it existed on a district and school level in the hopes that these tests would give a more individualized and accurate picture of the achievement of an individual student.

This process would be three-fold. First, a clear curricular procession for each subject being tested would need to be established, with the important curricular markers identified at each testable level. Next, the test items would be tied to the curriculum objective that it was meant to assess. In this way, each item on the test would relate directly back to a learned piece of the curriculum. Finally, the scoring of these tests would be done based on the level within the curriculum. This way, progress for a student could be charted in each individual tested area, increasing the relevance of the results of these tests to the students themselves (Cox & Sterrett, 1970).

Embedded in the debate over whether these tests were a good idea was the idea of the inherent bias in the test questions. Minority groups including African Americans and women felt that the questions were not fairly phrased to allow their groups to determine the correct answer and garner the same score as, say, a white male. In a paper written by Jane Faggen-Steckler, Karen McCarthy, and Carol Tittle in the *Journal of Educational Measurement* in 1974, eight standardized tests were examined for “sex bias.” The study found that all eight examined tests, including the California Achievement Tests, the Iowa Test of Basic Skills, the Stanford

Achievement Tests, and the Metropolitan Achievement Tests, exhibited bias based on sex (Faggen-Steckler, McCarthy, & Tittle, 1974).

The study examined names, pronouns, and nouns in order to determine the frequency at which each gender was used. Using a ratio, the researchers discovered that male nouns, names, and pronouns were used much more often in standardized tests. One test had a ratio of 14:1 male to female mentions, and all but one of the test included in the study had a ratio of over 1.5:1 (Faggen-Steckler, McCarthy, & Tittle, 1974). The other analysis included in the study had to do with gender roles and stereotyping. Again, the study found that females were placed in roles of the home; it was even noted that no adult woman mentioned in the tests that were examined had any occupation outside of the home. They determined that these imbalances were created by selections of the test item writers, rather than by a limitation of the English language or other source.

These bias questions and dissatisfaction with the process of testing can be attributed at least in part to the political turmoil that came into being in the late 1960s. As more groups advocated for their rights within and outside of education, concerns about equality to each group needed to be considered. This was not an easy period socially for the United States, and education was no different.

The decade saw massive changes in special education, including the court decision in *PARC v. Pennsylvania*, which determined that students with mental retardation also had a right to a free, public education. The definition of gifted and talented also changed during this period, and court-ordered desegregation of schools continued. Bilingual education became more accepted during this time period, and there was also a push for de-schooling and another for a move back to the basics of education (“American Educational History,” 2010). That any

institution from this period survived is fascinating, but that standardized testing came through the fog of the 1970s largely unchanged is miraculous. Many of the same arguments about standardized testing are still raised today, with the same proponents and detractors having the same fights from 40 years ago. That we still use them seems to indicate that the American education system sees some value in this method of assessment, but that the complaints and discussion are as heated as ever might indicate that the current method might not be the best.

The Impact on the Present Situation in Education:
From the Perspective of a Student and Teacher

So much of what happens in education comes from something that has been done decades ago and phased out, forgotten, and reintroduced as new. The influence and proliferation of standardized testing is no different. The National Defense Education Act (NDEA) of 1957 created a climate in which testing was the only conceivable way to determine the progress of students. As the technology caught up to the idea, testing exploded and an industry was created.

While the industry and tests never went away after that point, the fervor for testing faded, and the tests themselves came under scrutiny. Interest about testing and its importance ran high, and then the issue of standardized testing begins to fade from the literature. As it happens, the cycle continued and education continued to evolve. With the passage of the No Child Left Behind legislation in 2001, American education was essentially set on the same path from 44 years prior; testing was to become the paramount measure of educational success.

While there are several key differences between the years following NDEA's passage and NCLB's passage, many of the problems faced are shockingly similar. Both pieces of legislation essentially defined the focus of education, getting it back on track to improving achievement in key areas. NDEA did not, however, mandate any specific testing; NCLB demanded testing. This indicates a shift from a direct repetition of the cycle. Testing had been proven effective by

the intervening years; it was a necessary part of any large-scale program of educational improvement. These tests were to be chosen by the states to assess the standards that had been decided upon (The Elementary and Secondary Education Act (NCLB), 2001). The tests for this iteration of legislation were to be tied directly to a set of standards derived from research and curricula. This was definitely a new move.

Unfortunately, this move resulted in a similar result to the creation of NDEA. Testing became the only thing that mattered in schools in a much more rapid fashion than the research suggests happened in the late 50's and early 60s. As a student of the early NCLB era, I feel uniquely qualified to discuss the direct implications of this legislation on classroom instruction. It seemed like an immediate shift in English and mathematics classes. We were still learning the curricular material, yes, but we were also spending at least 1 class period a week preparing for the PSSA tests. It wasn't even a matter of reviewing skills that we should already have known, and in high school, that probably should have been the case, but we were instead learning about the format of the test, about the way it would be scored, and about how important this test was to the school. It had the effect on some students of making the test seem to be the most undesirable thing ever to happen in the already boring setting of school. Even the students who recognized the importance of doing well on the test were bored by the amount of preparation time that was being devoted to it. During my junior year of high school, parking permits and prom tickets began being offered as free rewards to students who scored proficient or above in all sections of the test. It was no longer at all about learning the material either in the classroom curriculum or in the context of these tests; it was about the rewards.

The same can be said for the school districts. With flexible pass rates and magical mathematical formulas for determining AYP, it wasn't the learning that mattered. It was the

numbers. Students who didn't score proficient were, in my school district, enrolled in a study skills course, which was, for all intents and purposes, a PSSA preparation class. While I understand the importance of these tests in determining the progress of districts toward offering good, research-based, proven educations to each student, the recipients of this extra attention associated more and more negatively with the test.

Now, as a student teacher, I can add a new layer to my understanding of this process. The textbook I use in my classroom, MacDougal Littell's Literature 9 Textbook, is written specifically to the PA standards, and all of the exercises included in the book are related directly to the format and scoring of the PSSAs. The standard of a good education is now only one thing: a proficient score on the PSSA test. With graduation tied to the PSSA requirements, it seems that this connection between this one test and the fate of an entire thirteen-year education will not dwindle.

But, as the education cycles continue to roll through, it might soon come into a new era similar to the 1970s. Testing lost its steam as a research interest, and companies began to review, revise, and renorm the tests they already had instead of continuing to write new tests all the time. The Iowa Test of Basic Skills, for example, is still the test of choice in Iowa, and its format has remained fairly constant throughout its entire usage ("ITBS," n.d.). The same format of teacher prompted directions is still an important part of the test, as can be seen in Figure 1. This part of the test is consistent throughout the versions of the test. The instruction reads the directions for the sections, which indicates that to the writers of the ITBS, this continuity of form from grades K-8 is important.

I conducted several short interviews with adults whose primary and secondary schooling took place during the 1960s and 1970s, and I learned more about the differences in the

implementation of these standardized tests in that time period. The overarching note that I got from my interviewees was that instructional time was not devoted to the preparation for any standardized test. Two of the three interviewees actually said, “They just gave it to you and said, ‘Take this’” (Pebley, 2010). The students were aware that the tests were, in large part, used to determine what levels of various subjects they could take the next year or semester, but all three indicated that there was no pressure from the teachers to perform perfectly on the assessments (Pebley, 2010).

When asked why they thought there is so much more pressure on students to test well now, all three answered the same way: “Money,” they said (Pebley, 2010). During the 1960s and 1970s, funding for schools was not directly tied to the results of test scores. The scores were determiners of student progress, but they were not the determiners of school and district progress. The scope of application has widened so significantly in the intervening 40 years that a school’s budget is effectively tied to the percentage of students who reach the level of proficiency on the state-determined standardized test.

One of the interviewees provided a particularly interesting insight into the change in the “mood” of testing from the time she was a student to now:

When I was in school, you were only in school for you. Your grades mattered to you and your parents; they determined *your* success and failure. That was all.

Now, we’re asking students in elementary school to bear the burden of essentially funding their teacher’s job for the next year. It’s a lot more stressful; it’s no wonder they don’t perform well and dislike the whole process (Pebley, 2010).

This mentality is one that explains, I think, the difference between the spirit of standardized testing and the reality of their use. Instead of using tests as indicators of an individual’s

achievement, which is really all one test can effectively measure, the culture of education has generalized the scope of testing data to encompass the determination of success or failure of an entire state's educational endeavors.

It is true that a great number of failing students indicates a problem with the education that those students receive, but tying funding and staffing concerns to the results of student tests creates an atmosphere of urgency and fear around testing that simply doesn't need to be there. Yes, they have a purpose. Unfortunately, that purpose in the twenty-first century seems to be to determine the entirety of worth of a school, district, or state. The test no longer tells us about the individual in any aspect except other than if they need remediation classes.

It is clear that the 1960s and 1970s have left a mark on education that has not and probably will not be erased. The proliferation of standardized testing continues today, as does the discussion surrounding the many issues within testing. This method of assessment has been a relevant topic of discussion through times of political and social turmoil in the country, and in this interesting political and social time, that relevance is more important than ever. The attention paid to standardized testing will hopefully lead to substantive decisions this time around, so that in 40 years, educators, researchers, and politicians need not be having the same discussions that are happening now.

References

- 1968: a timeline of events. (n.d.). *Scholarly Technology Group - Brown University Library*. Retrieved November 7, 2010, from <http://www.stg.brown.edu/projects/1968/reference/timeline.html>
- Airasian, P. (1987, May). State mandated testing and educational reform: Context and consequences. *American Journal of Education*, 95 (3), pp. 393-412.
- American Educational History Timeline. (2010). *Cloudnet*. Retrieved November 8, 2010, from <http://www.cloudnet.com/~edrbsass/educationhistorytimeline.html>
- Clark, D., & Guba, E. (Director) (1965, October). An examination of potential change roles in education. *Seminar on Innovation in Planning School Curricula*. Lecture conducted from National Education Association, Washington, D.C.
- Copperman, P. (1979, June). The achievement decline of the 1970s. *The Phi Delta Kappan*, 60 (10), pp. 736-739.
- Cox, R., & Sterrett, B. (1970). A Model for Increasing the Meaning of Standardized Test Scores. *Journal of Educational Measurement*, 7(4), 227-228.
- Faggen-Steckler, J., McCarthy, K., & Tittle, C. (1974). A Quantitative Method for Measuring Sex "Bias" in Standardized Tests. *Journal of Educational Measurement*, 11(3), 151-161.
- Flatteau, P. E., Bracken, J., Van Atta, R., Bandeh-Ahmadi, A., de la Cruz, R., & Sullivan, K. (2006). *The National Defense Education Act of 1958: Selected Outcomes*. Institute for Defense Analyses, Science and Technology Policy Institute. Institute for Defense Analyses.

- Glatthorn, A., Boschee, F., & Whitehead, B. (2009). *Curriculum leadership: Strategies for development and implementation* (second edition). Thousand Oaks, CA: Sage Publications, Inc.
- Houghton Mifflin Company -- Company History. (n.d.). *Connecting Angel Investors and Entrepreneurs*. Retrieved November 23, 2009, from <http://www.fundinguniverse.com/company-histories/Houghton-Mifflin-Company-Company-History.html>
- ITBS: Iowa Tests of Basic Skills - Iowa Testing Programs - College of Education - The University of Iowa. (n.d.). *College of Education - The University of Iowa*. Retrieved November 7, 2010, from <http://www.education.uiowa.edu/itp/itbs/Default.aspx>
- John Gardner - tackling the problems of urban America. (n.d.). *PBS: Public Broadcasting Service*. Retrieved November 7, 2010, from <http://www.pbs.org/johngardner/chapters/5.html>
- Karsten, E. (1971). Standardized Testing--Reality Therapy?. *The English Journal*, 60(8), 1062.
- Kelly, G. P., & Seller, M. S. (1985). *A history of school reform in New York state: Implications for today's policy makers*. Rockefeller Institute for Public Policy.
- Lemann, N. (1999). *The big test: the secret history of the American meritocracy*. New York: Farrar, Straus and Giroux.
- Linn, R. L. (1986). Educational testing and assessment: Research needs and policy issues. *American Psychologist*, 41 (10), 1153-1160.
- Marlaire, C., & Maynard, D. (1990). Standardized testing as interactional phenomenon. *Sociology of Education*, 63(2), 83-101.
- Pebley, E. (2010, November). [Interview with Christina T. Pebley].

Pebley, E. (2010, November). [Interview with Sonya Riden].

Pebley, E. (2010, October). [Interview with Stephen Pebley].

Riverside Publishing - Iowa Tests of Basic Skills (ITBS) details. (n.d.). *Riverside Publishing*.

Retrieved November 7, 2010, from <http://www.riverpub.com/products/itbs/details.html>

Rudman, H. (1977). The Standardized Test Flap. *The Phi Delta Kappan*, 59(3), 179-185.

Secrets of the SAT: Where did the test come from?: History of the SAT - a timeline. (n.d.). *PBS:*

Public Broadcasting Service. Retrieved November 7, 2010, from

<http://www.pbs.org/wgbh/pages/frontline/shows/sats/where/timeline.html>

The Elementary and Secondary Education Act (The No Child Left Behind Act of 2001). (2004,

September 14). *Ed.gov*. Retrieved November 7, 2010, from

www2.ed.gov/policy/elsec/leg/esea02/index.html

Thomas, J., & Brady, K. (2005). The Elementary and Secondary Education Act at 40: Equity,

Accountability, and the Evolving Federal Role in Public Education . *American*

Educational Research Association, 29, 51-67.

Urban, W. J. "The National Defense Education Act (NDEA): Much more than science or

Sputnik." *Midwest History of Education Society*. Chicago, IL. Oct. 2008.

Webster's encyclopedic unabridged dictionary of the English language (New deluxe ed.). (2001).

New York: Random House Value Publishing, Inc.

Winchester, I. (2006). Standardized testing and the classroom. *The Journal of Educational*

Thought, 40(2), 103-106. Retrieved November 7, 2010, from Education Module.

(Document ID: 1157380961)

UNIT 4 Language Mechanics

Lesson 4: Using Correct Capitalization

Directions: Darken the circle for the line that has a capitalization error. Darken the circle for *No mistakes* if there are no capitalization errors.

TRY THIS First, read the sentence or sentences. Then look at each line for words that should be capitalized and for words that should not be capitalized.

S1 A Jake's favorite day of the
 B week is Tuesday, that's the
 C day he has his piano lesson.
 D (No mistakes)

THINK IT THROUGH The correct answer is B. The word that's should be capitalized because it is the first word of a sentence.

STOP

1 A Last summer I wrote to a
 B magazine and asked for a pen pal.
 C Now I have a friend in nevada.
 D (No mistakes)

2 J Katie was happy to hear that
 K aunt Beverly had won a
 L prize at the school carnival.
 M (No mistakes)

3 A Every summer, millions of bats
 B come from Mexico to make
 C their home in Austin, Texas.
 D (No mistakes)

4 J For father's day, I gave
 K my dad a funny tie with a
 L picture of Groucho Marx on it.
 M (No mistakes)

5 A We drove down the Street
 B to the market so that we could
 C buy some bread, milk, and eggs.
 D (No mistakes)

6 J My favorite book by E.B. White is
 K charlotte's web. It is about a spider
 L who writes things in her web.
 M (No mistakes)

7 A Mark and Alex like to ride
 B their bicycles after school.
 C They go to roosevelt park.
 D (No mistakes)

8 J Saving money is important to me.
 K Each month i put three dollars
 L from my allowance into the bank.
 M (No mistakes)

GO ON

Answers S1: ● ○ ○ ○ ○ 2: ○ ● ○ ○ ○ 4: ● ○ ○ ○ ○ 6: ○ ● ○ ○ ○ 8: ○ ● ○ ○ ○

14 1: ○ ● ○ ○ ○ 3: ○ ○ ○ ● ○ 5: ● ○ ○ ○ ○ 7: ○ ● ○ ○ ○

24 Level 10

UNIT 4 Language Mechanics

Lesson 4: Using Correct Capitalization

Language Skills: Recognizing correct capitalization in sentences; identifying capitalization errors in sentences

SAY: Turn to Lesson 4, Using Correct Capitalization, on page 14.

Check to see that all students find Lesson 4.

SAY: In Lesson 4 you will practice finding capitalization errors in sentences.

Read the Directions to students.

SAY: Now look at Try This.

Read Try This to students.

SAY: Now look at S1. Read the lines silently. Then darken the circle for the line that has a capitalization error. Darken the circle for No mistakes if there are no capitalization errors.

Allow students time to choose and mark their answer.

SAY: Now look at Think It Through.

Read Think It Through to students. Check to see that all students have filled in the correct answer space. Ask students if they have any questions.

SAY: Now you will practice finding more capitalization errors in sentences. Put your finger on number 1. Do numbers 1 through 19 just as we did S1. When you come to the words GO ON at the bottom of page 14, continue working on the next page. When you come to the word STOP at the bottom of page 15, put your pencils down. You may now begin.

Allow students time to choose and mark their answers.

Figure 1. Image of a page from the teacher direction manual for the Iowa Test of Basic Skills.

(image courtesy of <http://mariinc.com>)

Figure 1. How a Standardized Test Is Constructed

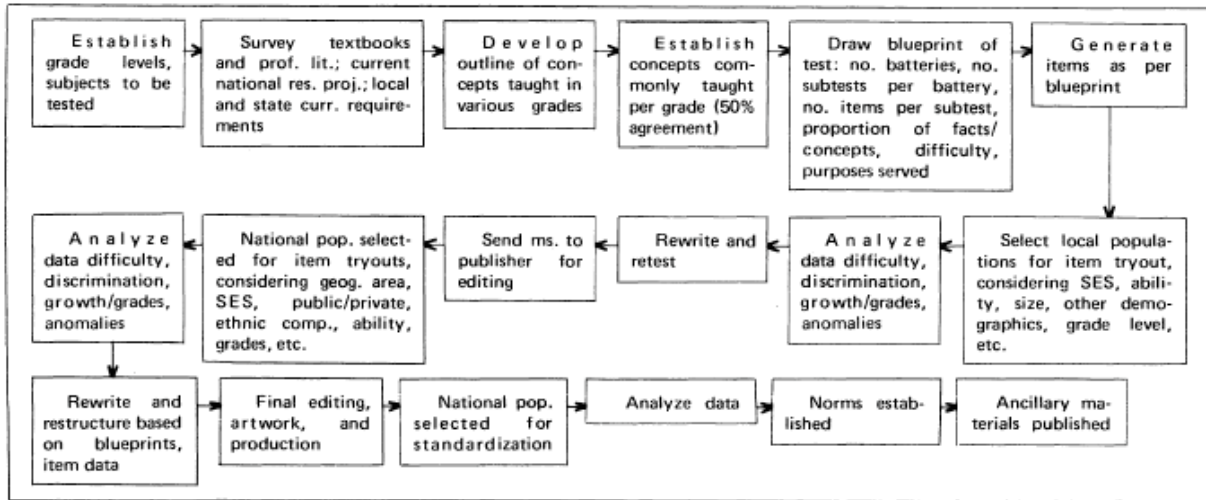


Figure 2. Image detailing the process of standardized testing construction from 1977 (image courtesy of “The Standardized Test Flap”)

Academic Vita of Emma E. Pebley

Name: Emma Elizabeth Pebley

Address: 111 Fairview Drive
Harrisburg, PA 17112

E-mail ID: eep5026@psu.edu

Education

Major and Minor: Secondary Education: English/Communications

Minor in English

Honors: Education

Thesis Title: Standardized testing in the 1960s and 1970s: Exploring the impact on today's educational climate

Thesis Supervisor: Dr. David Gamson

Work Experience

Date: Fall 2008-Fall 2010

Title: Undergraduate Research Assistant

Description: Photocopied, scanned, and catalogued elementary reading textbooks spanning from 1895 to 2008. Coded .txt files to make the files compatible with the analysis program designed for the project Researched and noted anomalies in textbooks and eras; Located additional textbooks or research texts when necessary.

Institution/Company: Spencer Reading Project (The Pennsylvania State University)

Supervisor's Name: Dr. David Gamson
Sarah Eckert