

# Quizzes, Tests, Exams, Assessments; Is There Any Time Left for Instruction?

Submitted for Completion of MS in Mathematics Education  
Buffalo State College

by Craig M. Dana  
September 27, 2001

# Table of Contents

<b><u>CHAPTER 1</u></b> .....	<b>4</b>
<b><u>Introduction</u></b> .....	<b>4</b>
<u>Increased Standards</u> .....	4
<u>Changes in Assessment</u> .....	5
<b><u>Government’s Influence on Testing</u></b> .....	<b>6</b>
<u>President Bush’s Agenda</u> .....	6
<u>Title I’s Influence on Testing</u> .....	6
<u>New York State’s Assessment System</u> .....	7
<b><u>Conclusion</u></b> .....	<b>7</b>
<b><u>CHAPTER 2</u></b> .....	<b>9</b>
<b><u>Introduction</u></b> .....	<b>9</b>
<b><u>Types of Standardized Tests</u></b> .....	<b>9</b>
<u>Differences in the Types of Answers Required</u> .....	9
<u>Multiple-Choice Assessments</u> .....	9
<u>Performance Based Assessments</u> .....	10
<u>Differences in The Use of the Results</u> .....	10
<u>Criterion-referenced Tests</u> .....	11
<u>Norm-referenced Tests</u> .....	11
<b><u>Arguments For Standardized Testing</u></b> .....	<b>11</b>
<u>Improve Low Scores</u> .....	12
<u>Identify Strengths and Weaknesses</u> .....	12
<u>For Students</u> .....	12
<u>For Teachers and Schools</u> .....	13
<u>Increase Motivation</u> .....	13
<u>For Students</u> .....	13
<u>For Teachers and Schools</u> .....	14
<b><u>Arguments Against Standardized Testing</u></b> .....	<b>14</b>
<u>High Stakes Negatively Affect Instruction</u> .....	15
<u>Cost</u> .....	16
<u>The Tests are Invalid Measures of Achievement</u> .....	16
<u>Testing Takes Time Away from Instruction</u> .....	18
<b><u>CHAPTER 3</u></b> .....	<b>20</b>
<b><u>Introduction</u></b> .....	<b>20</b>
<u>The Purpose</u> .....	20
<u>The Process</u> .....	20
<b><u>The Survey</u></b> .....	<b>20</b>
<u>The Start of the Survey</u> .....	20
<u>Types of assessment</u> .....	21

<a href="#">Time spent reviewing, testing and correcting</a>	22
<a href="#">Other interruptions</a>	23
<a href="#">Additional comments</a>	23
<b><a href="#">Conclusion</a></b>	<b>24</b>
<b><a href="#">CHAPTER 4</a></b>	<b>25</b>
<b><a href="#">Introduction</a></b>	<b>25</b>
<b><a href="#">The Results</a></b>	<b>25</b>
<a href="#">Overall Findings</a>	25
<a href="#">Effects in Specific areas</a>	28
<a href="#">Types of Assessment</a>	28
<a href="#">Time Spent Reviewing</a>	29
<a href="#">High School vs. Middle School</a>	29
<a href="#">Urban vs. Suburban Schools</a>	30
<a href="#">Teacher Comments</a>	30
<b><a href="#">Conclusion</a></b>	<b>33</b>
<b><a href="#">APPENDIX A FORM</a></b>	<b>35</b>
<b><a href="#">APPENDIX B DATA</a></b>	<b>43</b>
<b><a href="#">References</a></b>	<b>49</b>

## CHAPTER 1 Introduction

### Increased Standards

There are two main issues in education today, a call for increased standards and more accountability through assessment. For mathematics teachers, an increase in standards is not a new idea. The National Council of Teachers of Mathematics (NCTM) came out with the Curriculum and Evaluation Standards for School Mathematics in 1989, the Professional Standards for Teaching Mathematics in 1991, and their newest set of standards, Standards 2000, just last year.

Prior to increased standards, math teachers generally followed a text to establish their curriculum, and might supplement a few activities. In general, it was clear what was expected and pretty much everyone taught that way. (Alper, 1995, p. 598) The NCTM Standards have changed this philosophy. Developing students' mathematical power is one of the main focuses of NCTM's Professional Standards. In order to meet these standards, teachers must adopt different teaching techniques, as compared to those they used to use. These new techniques are defined as building a variety of tools to allow students to solve problems, as well as allowing students to solve problems in various ways, and not discouraging them from using "less elegant approaches." (Alper, 1995, p. 601) Another technique that fosters students' mathematical power is structuring good mathematical tasks. Utilizing good mathematical tasks means that the students may not immediately know how to solve a problem, but the teacher allows them to struggle and take responsibility for the learning. (Smith, 2000, p. 379)

In comparison to the traditional style of teaching, these techniques may require a lesson to take more time. For example, a lesson on dividing fractions ten years ago might have been done in 20 minutes, with 30 examples for the students to complete for homework. In order for

the students to understand what it means to divide fractions, they need to struggle, question, and discuss a problem that will lead them to that understanding. The new ideas and methods outlined by NCTM provide opportunities for all students to learn and understand mathematics. For this reason, most mathematics teachers have embraced the ideas, strategies and techniques outlined in NCTM's Standards. However, increases in assessment have not been as well received.

### Changes in Assessment

It seems that with low achievement on national and international tests, there are higher expectations and an increase in the number of standardized tests in America. By 1980, there were 38 states that had state testing programs. In a survey completed in 1997, the number had increased to 46 states. (Mehrens, 1998) The increase in testing, compared to the increase in standards, is a much more strongly debated topic. Gregory Anrig (1992) writes,

“Many businesspeople and elected officials love tests; they see them as the end-all and be-all that will solve all of the problems in education. On the opposite side, many educators, and particularly teachers, see the growth in accountability testing as an unqualified evil and call for the abolishment of standardized testing.” (p. 3)

As Anrig states, there are definitely two sides to this issue. Mainly the politicians and businesspeople are encouraging the increase in testing, while many educators are opposing the increase.

With new standards that teachers are required to meet and new assessments that students have to take, when is there time to teach? Implementing new teaching techniques that require more time can be difficult when there are more tests taking time away. Who is requiring the increase in assessments, and why?

## Government's Influence on Testing

### President Bush's Agenda

The main focus of President Bush's Educational agenda is annual testing in grades 3 through 8. This program is modeled after the state testing system while he was the Governor of Texas. The Texas Assessment of Academic Skills is a criterion-referenced test designed to measure students' achievement in mathematics and reading in grades 3 through 8, and at the secondary level. In some of those grades, the Texas test also measures writing, science, and social studies. Texas looks at each grade level to see if students have mastered certain concepts, and schools are held accountable for minimum passing rates for all students. (Olson, 2001)

While 46 states currently have testing programs, most of these assessments do not resemble Texas's testing system. This may cause problems when President Bush begins to implement his program nationwide. The President's answer to this dilemma is federal aid. President Bush may propose \$100 million of federal aid to assist states in writing and administering tests in the first couple of years. Also, while Bush's plan encourages states to use standards based tests, there is the option to use a less expensive norm-referenced test if needed. (Olson, 2001)

### Title I's Influence on Testing

In 1994 Congress rewrote the Chapter 1 law for disadvantaged students, naming the new law Title I. Chapter 1 required that districts track the academic achievement of the students in this program in grades 2 through 12, using a norm-referenced test. Title I now requires that states and districts use criterion-referenced tests that actually reflect their state's standards. Also, these tests should be the same tests used to measure the performance of other students in the state. These tests are required to be given once annually in grades 3 through 5, 6 through 9, and

10 through 12. (Olson, 2001) Title I has been a major factor in states moving from multiple choice, norm-referenced tests, to short answer, criterion-referenced tests.

### New York State's Assessment System

New York State currently tests students in 4<sup>th</sup> grade, 8<sup>th</sup> grade, and at the secondary level. These levels meet the guidelines established by Congress, because there are tests between 3<sup>rd</sup> and 5<sup>th</sup> grade, 6<sup>th</sup> and 9<sup>th</sup> grade, and 10<sup>th</sup> and 12<sup>th</sup> grade. New York State uses the Intermediate tests in 4<sup>th</sup> and 8<sup>th</sup> grade to measure how well students are meeting the state's standards, and the high school assessments are used as requirements for graduation. In the 2000-2001 school year, 4<sup>th</sup> graders will have taken assessments in English Language Arts, Mathematics, and Science. The eighth-graders, in May and June of 2001, will have taken a total of 7 tests in 6 content areas. These tests include English Language Arts, Mathematics, the Science written test, the Science performance test, Social Studies, Technology, and Foreign Language.

New York State uses these tests to monitor students, and hold schools accountable. Each school is given a report card and is held to certain standards. Often times the results of these tests are published in local papers, comparing how schools perform locally and across the state.

### Conclusion

Clearly, there exists a great deal of mandated testing in schools, not including the classroom assessment the teachers do themselves. Therefore, with increased standards and testing, how much time is left for instruction? This study will focus on how much time is actually devoted to testing in our schools? We will look at required standardized assessments, as well as regular classroom assessment. Logically, with an increase in testing there should also be an increase in reviewing. If this is true, are students spending most of their time preparing, taking, and reviewing tests, or is most of their time spent learning? Does the presence and

pressure of upcoming tests effect the method in which teachers instruct their students? This study will address each of these questions in the upcoming chapters.



## CHAPTER 2 Introduction

There are two debates on standardized testing, whether or not standardized tests improve education, and whether or not standardized tests are reliable measures for student achievement. This chapter will analyze a portion of the research regarding these two topics. First we will discuss the types of standardized tests. Then we will look at the different reasons for giving and for not giving standardized tests. This will help us understand the entire issue before we begin to look, in the last two chapters, at how assessments effect instructional time.

### Types of Standardized Tests

#### Differences in the Types of Answers Required

##### Multiple-Choice Assessments

Multiple-choice tests, which have been popular forms of assessment for decades, require students to choose the best answer from a given list of choices. A couple of reasons for their wide use are that they are easily graded, as well as they are relatively inexpensive to administer. Multiple-choice tests do not require space for students to show their work, thus lowering the cost of the test booklets. Also, with the use of bubble sheets, a computer can grade multiple-choice tests very quickly and efficiently and give teachers immediate feedback on which items were answered correctly and which items were answered incorrectly. The abundant use or misuse of multiple-choice tests has lead to some recent criticism. For example, David Berliner (1992, p. 11) writes in his article supporting performance based testing, that Americans have finally seen the limits of multiple-choice standardized tests and how these tests negatively influence instruction.

## Performance Based Assessments

A performance based test asks the students to write an open response. Many educators believe that performance based testing allows for students to better demonstrate their knowledge, and for teachers to better grade students' understanding. Performance based tests get closer to the skills that teachers are really looking for. (Berliner, 1992, p. 11) One drawback of a performance based test is the cost. Performance based tests are much more expensive to administer than multiple-choice tests. This is due to the actual cost of the test booklets, in which there needs to be enough space for students to show their work, as well as the cost needed to grade the tests. Performance based tests need to be graded by a teacher, which means that school districts must either pay teachers or companies to score their tests. The district where I teach is going to pay 8 or 9 teachers \$25 per hour to grade the State's Social Studies Assessment in June. To grade the ELA, Math and Science tests, 8 or 9 teachers from each department will be out of the classroom for two days. This requires substitute teachers to be paid, while these 8 or 9 teachers, per department, per day, score the tests. When multiplying that amount by the total number of schools across the state, and the total number of assessments, English, Math, Science, Technology and so on, performance based tests can become extremely expensive to grade.

### Differences in The Use of the Results.

How the test results are analyzed is the difference between a criterion-referenced test and a norm-referenced test. Both tests can be either multiple-choice or performance based, however norm-referenced tests generally tend to be associated with multiple-choice tests, while criterion-referenced tests are generally associated with performance based tests.

### Criterion-referenced Tests.

A Criterion-referenced test is designed specifically to measure students' progress with respect to a particular state's standards. Many states use criterion-referenced tests to provide information not only on the students, but also to provide information on how well school districts are meeting their standards. New York State is currently giving performance based, criterion-referenced tests in 4<sup>th</sup> grade, 8<sup>th</sup> grade, and in high school. New York uses these results to see if students and schools are meeting the State's standards.

### Norm-referenced Tests.

A norm-referenced test uses its results to rank students in comparison to their peers from around the country. An example of a norm-referenced test, that many schools are currently using, is the Terra Nova, published by CTB/McGraw-Hill. A student who has taken the Terra Nova will get their score as a percentile. The results of a norm-referenced test provide such information as to how many students scored above or below you, where the average score was, and where the lowest and highest scores fall. These types of tests are often used to assist in identifying students who need special attention. We have looked at the different types of standardized tests, next we will explore the reasons in favor and against using standardized tests.

### Arguments For Standardized Testing

The politicians and businesspeople that favor standardized testing have some valid reasons for implementing assessment programs. First and foremost, the government and business world want American students to score at the top of the list on national and international tests. When students have low scores, there tends to be a push for more assessments designed to improve achievement. Another reason in favor of testing is that they provide valuable

information on students, teachers, and schools. Also, when high stakes are attached to the tests, they can be used to motivate students and schools to improve achievement.

### Improve Low Scores

The government is one large group in favor of standardized testing. As noted earlier, President Bush, as well as many state governments, are pushing for more standardized tests. In President Bush's plan, No Child Left Behind, he outlines some of the reasons for an increase in testing. President Bush (2001) writes,

“If our country fails in its responsibility to educate every child, we're likely to fail in many other areas. But if we succeed in educating our youth, many other successes will follow throughout our country and in the lives of our citizens. (Foreword)

President Bush is responding to the feeling that our nation is falling behind in comparison to other nations. President Bush (2001, p. 1) continues by stating that 70% of inner city 4<sup>th</sup> graders are unable to read at a basic level, and that our high school seniors are below students in Cyprus and South Africa on international math tests. Generally, educational policies are mandated by local school boards, but with decreasing national and international test scores, the government feels that there is a national crisis, giving them reasons to get more involved.

### Identify Strengths and Weaknesses

#### For Students

Politicians are not the only ones in favor of standardized testing. Many educators believe that standardized tests provide valuable information when used properly. Many of those educators feel that when students get the results back from a test, the scores can identify their strengths and weaknesses. For example, a student could score very well in mathematics and science, but score poorly in reading. Teachers and parents can then focus on the areas where

students are weak, and thus increase overall achievement. This also applies in a specific content area. With respect to mathematics, a test may show that a student does well in algebra and computation, but scores low in geometry. Again, this provides an opportunity to assist the student. According to James Popham (1999, p. 21), standardized achievement tests do a wonderful job of supplying the evidence needed to make norm-referenced interpretations of students' knowledge and/or skills in relationship to those of students nationally. Also, these "norm-referenced interpretations of students knowledge" provide schools with information to rank and group students. Standardized tests have assisted teachers, counselors, and administrators for many years in identifying students for remedial and advanced classes.

#### For Teachers and Schools

Other than identifying strengths and weaknesses for students, test results can be used to track the progress of teachers and schools. Rosaena Garza, the director of academics for a school district in Texas believes that annual testing is a powerful tool for principals. For example, a principal can look for either positive or negative patterns in a particular teacher's classroom, and then begin to investigate what is happening. (Olson, 2001) This allows teachers to adopt techniques from teachers who have had an increase in student achievement, which can work to benefit the students of an entire school.

#### Increase Motivation

##### For Students

Standardized testing can be an intense motivator for students when there are high stakes attached to the tests. Many states use assessments in the high school as part of a graduation requirement. For example, in order to graduate in New York State a student must pass the Math A exam. This form of high stakes testing is designed to demonstrate to students the importance

of academic standards. If a student does not meet New York State's expectations, he or she will not receive a diploma. High Stakes testing may motivate the students to improve their academic achievement, but can also have an effect on teachers and schools.

### For Teachers and Schools

With appropriate incentives and/or sanctions, assessments can motivate students to learn better, teachers to teach better, and schools to be more effective. (Linn and Herman, 1997, p. iii)

As outlined by President's Bush plan, motivation is an important topic. Schools that meet President Bush's expectations will be rewarded, and those who do not meet the President's expectations, may lose their funding. Providing and taking away funds can play a large part in motivating schools to improve academic achievement. Also, schools' test scores are often published in newspapers for public criticism or praise. New York State provides parents with a school report card. Schools are motivated to increase scores, not only to keep funding, but also to insure that their school has a favorable public opinion. Not all educators are convinced that attaching high stakes to tests motivate schools and students for the right reasons. This leads us to look at some opposing views to standardized testing.

### Arguments Against Standardized Testing

There are many reasons why teachers, parents and students might be against standardized testing. David Berliner (1992) writes,

“It is sad, but our politicians and many citizens consider students to be of little value unless they are certified “smart” by tests in the areas that have been chosen by society to be emphasized.” (p. 13)

In a recent poll, 46% of teachers said there is too much emphasis on testing, while only 20% of the public had the same view. Also, 10% of teachers say that there is too little emphasis on

testing, while 28% of the public wanted an increase. A majority of teachers, 69%, were opposed to then President Clinton's proposal for voluntary national assessments, while a majority, 57%, of the public was in favor of his plan. (Langdon, 1997, p. 213)

There are many reasons why educators are opposed to standardized testing. Many educators feel that high stakes testing has a negative effect on instruction. Also, the expenses required by a state to implement a testing program can be astronomical. Another reason is that the tests themselves might be invalid and/or not measure what they intend to. And lastly, and most importantly, too much assessment negatively effects the amount of time available for instruction.

#### High Stakes Negatively Affect Instruction

As with anything, when money becomes attached to a situation, it can have a negative effect. For example, when students are playing a game, there is a different attitude when money is involved and when the game is just for fun. When high stakes become a part of the situation, an attitude of doing anything to improve can prevail. In the case of testing, schools and teachers mobilize to do well on the tests. The teachers often abandon the types of activities needed to foster mathematical power. The program becomes one of drill and practice covering only the topics on the tests. This then results in a narrowing of the curriculum. (Berliner, 1992, p. 12) (O'Brien, 1989, p. 361) (Romberg & Wilson, 1992, p. 18) Also, high stakes testing has effected other subjects and activities. Schools across the country, under the pressure to score well on tests, have cut back such programs as recess for young children, classes in the arts, electives at the high school, class meetings and other activities to promote social and moral learning. (Kohn, 2001, 250)

## Cost

When administering a standardized assessment, the amount of cost can be the first problem a state faces. In 1990, a Ford Foundation-funded study, estimated that the direct costs of testing in schools is about \$100 million a year. (Anrig, 1992, p. 3) Achieve, a non-profit group that promotes standards based initiatives, reported that in the past five years state testing costs have gone from \$141 million to \$390 million. This is an increase of almost 400% from the Ford Foundation survey in 1992 to the Achieve survey in 2001. More specifically, the estimated average cost per student for a multiple-choice test is about \$17.50, and a performance based test costs an average of about \$28 per student. (Olson, 2001) If a state needs \$28 per student in order to implement its assessment program, where is that money going to come from, or what could they do if they did not spend that money on testing? Schools could be using that money to purchase new materials, technology, funding for teacher in-service, and to fill many other educational needs.

## The Tests are Invalid Measures of Achievement

Susan Ohanian (2001) in her article, “News From the Test Resistance Trail,” discusses different types of questions from some standardized tests. She writes,

“Fourth graders in New York City are interrogated about the purity of maple syrup; high-schoolers are asked to respond to an essay by Roger Ascham. Students in Los Angeles are asked about lemon mousse. The SAT 9 (Stanford Achievement Test) gives third graders a rigorous proofreading test, along with some nutty vocabulary terms. More important, does anyone think the answers should determine whether a student passes or fails third grade?” (p. 363)



Ohanian is not alone in feeling that many test questions are foolish and do not really show how much a student knows. In another example, 15 out of 20 students knew the date of Jefferson's embargo, but only one student knew what "embargo" meant. (Bracey, 1997, p. 86) (O'Brien, 1989, p. 361) Not only have the questions received criticism, but how the process by which tests are administered have been called unfair. Ohanian (2001, p. 363) also points out that students taking the Graduate Record Examination or the Law School Admissions Test are given more time per item than is given to students taking the SAT 9. Not only have the questions and procedures of standardized testing been under attack, but their results have also been denounced.

Alfie Kohn (2001, p. 349) believes that standardized tests, in part, only measure socioeconomic status. He states that socioeconomic status accounts for an overwhelming proportion of variance in standardized testing. If your school mostly has students who come from a genuinely low socioeconomic situation, and your students do not have high scores, does that mean your teachers are ineffective? On the other hand, if your students are upper class and have well-educated parents, when they score well, does that mean your teachers are better? Of course not. The issues that your students bring to school from their out of school environment have just as much, if not more affect on their academic achievement. (Popham, 1999, p. 7) All of these reasons against testing and all of the reasons for testing can ardently be debated.

However, if teachers are spending less time with their students, education will suffer. Research has shown that students who spend more time on task, score better. This only makes sense. If one class spends 20 weeks on a topic and another class spends 30 weeks, the second class will score better. This is why the focus of this study is on time. In the next section we will look at how much time is actually needed for mandated standardized tests.

### Testing Takes Time Away from Instruction

Now that we understand the different types of standardized tests that are used, as well as the arguments for and against testing, we can begin to look at the effects on instructional time. The first direct effect of standardized testing is that too much class time is spent administering the tests. Laurie B. Abeel, the assessment coordinator for a Virginia school district says, "I'm testing all the time, when do I have time to teach?" (Olson, 2001) Peter Simon's (2001, A1) article, "Testy time for schools," takes an in-depth look at the number of assessments that a typical eighth-grade student will take in the Buffalo City School District. The students will spend a sizable part of 18 schooldays practicing for, and taking state assessments. When compared to the 180-day school year, that is 10% of the entire year spent just on assessment. Also, eighteen school days is almost 4 school weeks, which is close to half of an entire marking period spent not teaching, yet the assessment does not stop there.

Basically, we are spending too much time assessing students and not enough time teaching them. The 18 days mentioned in the Buffalo News article were the number of days required for testing by each school. The article does not account for all of the days individual teachers spend reviewing and testing prior to the tests. The Ford Foundation-funded study mentioned earlier reported that nearly 20 million student days per year are spent on testing. (Anrig, 1992, p. 3) That was in 1992, just imagine how the total number of student days spent testing has risen with the increase of more testing.

There are many articles regarding standardized testing, however, there is not much research investigating general assessment practices with respect to instructional time. One example of how much time can be spent on testing is a teacher who spent one fourth (6 days out of 23 days) in a variables and patterns unit on assessment. (Keiser & Lambdin, 1996, p.28)

According to Stiggins (1988, p. 363), the everyday assessments used to determine class pacing and student grades have been ignored. Also, he states that teachers may spend as much as 20% or 30% of their professional time directly involved in assessment-related activities. This would include the time spent making, administering, scoring, recording and reporting such items as daily homework, tests, quizzes, and observations and judgments about student performance. This research was done over a decade ago. With the increases in testing, I wonder if these figures have risen. That will be the focus in Chapter 3. We will survey teachers to question how much time is actually spent assessing students.